

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB	ID	Description
1	1174	100.0	1174	6	A92311	A92311 Sequence 1
2	1174	100.0	1174	6	AR075389	AR075389 Sequence
3	1174	100.0	1174	6	AR094310	AR094310 Sequence
4	451.6	38.5	922	6	A68070	A68070 Sequence 65
5	451.6	38.5	922	6	AR163167	AR163167 Sequence
6	437.4	37.3	1060	6	A21792	A21793 H. insolens
7	437.4	37.3	1060	6	A23635	A23635 H. insolens
8	437.4	37.3	1060	6	A23644	A23644 H. insolens
9	437.4	37.3	1060	6	A23953	A23953 Endoglucanase
10	437.4	37.3	1060	6	A23957	A23957 Endoglucanase
11	437.4	37.3	1060	6	A41658	A41658 Sequence 1
12	437.4	37.3	1060	6	A68060	A68060 Sequence 55
13	437.4	37.3	1060	6	AR059002	AR059002 Sequence
14	437.4	37.3	1060	6	AR072921	AR072921 Sequence
15	437.4	37.3	1060	6	AR163160	AR163160 Sequence
16	437.4	37.3	1060	6	I13884	I13884 Sequence 1
17	437.4	37.3	1060	6	I23136	I23136 Sequence 1
18	437.4	37.3	1060	6	I57983	I57983 Sequence 1
19	437.4	37.3	1060	6	AR220054	AR220054 Sequence
20	437.4	37.3	1060	6	AR261282	AR261282 Sequence
21	437.4	37.3	1060	6	BD010852	BD010852 Cellulase
22	438.6	37.3	1060	6	BD015226	BD015226 Non-toxic
23	428.6	36.5	1057	6	BD002248	BD002248 Cellulase
24	411.4	35.0	919	6	A68072	A68072 Sequence 67
25	411.4	35.0	919	6	AR163168	AR163168 Sequence
26	393.4	33.5	155568	6	NCH19917	BX284763 Neurospor
27	392.2	33.4	807	6	AR261284	AR261284 Sequence
28	392	33.4	922	6	A68078	A68078 Sequence 73
29	392	33.4	922	6	AR163171	AR163171 Sequence
30	377.2	32.1	1257	6	AR121354	AR121354 Sequence
31	377.2	32.1	1257	6	BD144211	BD144211 Polymer-C
32	377.2	32.1	1257	6	BD168986	BD168986 Cellulase
33	377.2	32.1	1257	6	BD170331	BD170331 Zymocycl
34	375.6	31.9	1257	6	AR213263	AR213263 Sequence
35	374.6	31.9	1257	6	AB003107	AB003107 Humicola
36	369.8	31.5	1154	6	AR094308	AR094308 Sequence
37	364.6	31.1	924	6	A68074	A68074 Sequence 63
38	364.6	31.1	924	6	AR163169	AR163169 Sequence
39	353.6	30.1	936	8	MA151703	AJ151703 Melanocat
40	337	28.7	894	6	AR094306	AR094306 Sequence
41	336.2	28.6	912	6	A68076	A68076 Sequence 71
42	336.2	28.5	912	6	AR163170	AR163170 Sequence
43	334.8	28.5	927	6	AR094307	AR094307 Sequence
44	335.2	26.8	1443	8	F50KCH	L29381 Eubacterium ox
45	315.2	26.8	1473	6	A21795	A21795 F. oxysporum

JOURNAL Patent: EP 0843041-A 1 20-MAY-1998;  
NOVONORDISK AS (DK)  
FEATURES Location/Qualifiers  
source 1. 1174  
/organism="Saccharomyces cerevisiae"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4932"

## ORIGIN

Query Match 100.0%; Score 1174; DB 6; Length 1174;  
Best Local Similarity 100.0%; Pred. No. 3.2e-199;  
Matches 1174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
1 GAGGAGCAGCCCTCAAGCTGTAAGTTTCACCCCGCTCTTTCTTGGGCCCAAGA 60
1 GAGGAGCAGCCCTCAAGCTGTAAGTTTCACCCCGCTCTTTCTTGGGCCCAAGA 60
61 TGGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGGTGGCTCG 120
61 TGGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGGTGGCTCG 120
121 CGGCGAGTGGCAGTGGCCAGTCCAGAGATAGTGGGACCTGGTCAAGCCGCTT 180
121 CGGCGAGTGGCAGTGGCCAGTCCAGAGATAGTGGGACCTGGTCAAGCCGCTT 180
181 GAGCCGGGAAAGCGCGCTCAAGCCACCGGCTCAAGCGTGGATGCAATTCAGGCGC 240
181 GAGCCGGGAAAGCGCGCTCAAGCCACCGGCTCAAGCGTGGATGCAATTCAGGCGC 240
241 TGTTCGACTTCATATGTCAGTGGCGTGGCAAGCGTGGCGCTGCTGCTGCGCGAC 300
241 TGTTCGACTTCATATGTCAGTGGCGTGGCAAGCGTGGCGCTGCTGCTGCGCGAC 300
301 AGACTCCCTGGCGGCTGGAACGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
301 AGACTCCCTGGCGGCTGGAACGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
361 GCGGCTCCGAAATCTCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
361 GCGGCTCCGAAATCTCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
421 TGGCGGAGAGACATATGTTGTTGTCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 480
421 TGGCGGAGAGACATATGTTGTTGTCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 480
481 AGTTGATATGTCATATGTCATATGTCATATGTCATATGTCATATGTCATATGTC 540
481 AGTTGATATGTCATATGTCATATGTCATATGTCATATGTCATATGTCATATGTC 540
541 TGGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGGTGGCTCG 600
541 TGGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGGTGGCTCG 600
601 TGGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGGTGGCTCG 660
601 TGGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGGTGGCTCG 660
661 ACCCGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 720
661 ACCCGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 720
721 GCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
721 GCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
781 GCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
781 GCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
841 GTGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGGTGGCTCG 900
841 GTGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGGTGGCTCG 900
```

```
901 CTTGCTCTTGGGACCAACCTGCGCAAGATTGAACGATCTAATCTGAGTGGCTTAA 960
901 CTTGCTCTTGGGACCAACCTGCGCAAGATTGAACGATCTAATCTGAGTGGCTTAA 960
961 CAGCTTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
961 CAGCTTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
1021 TTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1021 TTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1081 TCTCAGACCGGCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
1081 TCTCAGACCGGCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174
1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174
```

RESULT 2  
AR075389 1174 bp DNA linear PAT 30-AUG-2000  
LOCUS  
DEFINITION  
ACCESSION AR075389  
VERSION AR075389.1 GI:10002139  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1174)  
AUTHORS Lund, H. and Kallum, L.  
TITLE Germline with considerable variation in abrasion level  
JOURNAL Patent: US 5958082-A 1 28-SEP-1999;  
FEATURES Location/Qualifiers  
source 1. 1174  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 1174; DB 6; Length 1174;  
Best Local Similarity 100.0%; Pred. No. 3.2e-199;  
Matches 1174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
1 GAGGAGCAGCCCTCAAGCTGTAAGTTTCACCCCGCTCTTTCTTGGGCCCAAGA 60
1 GAGGAGCAGCCCTCAAGCTGTAAGTTTCACCCCGCTCTTTCTTGGGCCCAAGA 60
61 TGGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGGTGGCTCG 120
61 TGGGCTCTAATCCCGTCTTGGCAACCCCTGGCGCTGCACTTCTCTGGTGGCTCG 120
121 CGGCGAGTGGCAGTGGCCAGTCCAGAGATAGTGGGACCTGGTCAAGCCGCTT 180
121 CGGCGAGTGGCAGTGGCCAGTCCAGAGATAGTGGGACCTGGTCAAGCCGCTT 180
181 GAGCCGGGAAAGCGCGCTCAAGCCACCGGCTCAAGCGTGGATGCAATTCAGGCGC 240
181 GAGCCGGGAAAGCGCGCTCAAGCCACCGGCTCAAGCGTGGATGCAATTCAGGCGC 240
241 TGTTCGACTTCATATGTCAGTGGCGTGGCAAGCGTGGCGCTGCTGCTGCGCGAC 300
241 TGTTCGACTTCATATGTCAGTGGCGTGGCAAGCGTGGCGCTGCTGCTGCGCGAC 300
301 AGACTCCCTGGCGGCTGGAACGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
301 AGACTCCCTGGCGGCTGGAACGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
361 GCGGCTCCGAAATCTCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
361 GCGGCTCCGAAATCTCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
```

QY 421 TCGCGCGCAAGCAATGCTGTGCTGACGACGACCTGCGCGACCTGCGAAGTAC 480  
DB 421 TCGCGCGCAAGCAATGCTGTGCTGACGACGACCTGCGCGACCTGCGAAGTAC 480  
QY 481 AGTTGATATGCGCATGCTCGCGCGCGCGTGGCATCTTCAACGCGCTGACGCTCGAGT 540  
DB 481 AGTTGATATGCGCATGCTCGCGCGCGCGTGGCATCTTCAACGCGCTGACGCTCGAGT 540  
QY 541 TCGCGCGCTCGCGCGCGCTCAATACGCGCGCATTTGTGCGCGACGACGATTCCT 600  
DB 541 TCGCGCGCTCGCGCGCGCTCAATACGCGCGCATTTGTGCGCGACGACGATTCCT 600  
QY 601 TCGCGCGCGCTCGACGCGCGCTGCGCGCGTGGCATCTTCAACGCGCGTGA 660  
DB 601 TCGCGCGCGCTCGACGCGCGCTGCGCGCGTGGCATCTTCAACGCGCGTGA 660  
QY 661 ACCGACGTTACGTTGACGACGATGACGCGCGCGCGATGCGTTCGCGCTCG 720  
DB 661 ACCGACGTTACGTTGACGACGATGACGCGCGCGCGATGCGTTCGCGCTCG 720  
QY 721 GCAAGCGCAACGACGACCTGCGCGCTTCAACCGCGCGACGCGTGG 780  
DB 721 GCAAGCGCAACGACGACCTGCGCGCTTCAACCGCGCGACGCGTGG 780  
QY 781 GCAAGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840  
DB 781 GCAAGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840  
QY 841 GTGGCTGACGCTGACGATGCGCGCTGACGCGCGCGCGCGCGCGCGCGCG 900  
DB 841 GTGGCTGACGCTGACGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
QY 901 CCGTGTCTCTGCGCAACGCTGCGAAGTTGAACGCTACTGCGAGTGCCTTAA 960  
DB 901 CCGTGTCTCTGCGCAACGCTGCGAAGTTGAACGCTACTGCGAGTGCCTTAA 960  
QY 961 CAGCTTTGCGACGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020  
DB 961 CAGCTTTGCGACGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020  
QY 1021 TTTTGAAGCGCTCAATACATACATAACCTTGTGATCTTGTACATAGACGCGTAC 1080  
DB 1021 TTTTGAAGCGCTCAATACATACATAACCTTGTGATCTTGTACATAGACGCGTAC 1080  
QY 1081 TCTCAACCGACTTTGGCGCGCGCGATGAGCGCGCTTTAA 1140  
DB 1081 TCTCAACCGACTTTGGCGCGCGCGATGAGCGCGCTTTAA 1140  
QY 1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174  
DB 1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174

RESULT 3  
LOCUS AR094310 1174 bp DNA linear PAT 08-SEP-2000  
DEFINITION Sequence 11 from patent US 6001639.  
ACCESSION AR094310  
VERSION AR094310.1 GI:10021119  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1174)  
AUTHORS Schulten,M., Andersen,L., Nornhoe., Lassen,S., Laasbedren,F.,  
Kauppinen,M., Sakari., Lange,L., Nielsen,R., Iium., Ihara,M. and  
Takagi,S.  
TITLE Endoglucanases  
JOURNAL Patent: US 6001639-A, 11-14-DEC-1993;  
FEATURES  
1. 1174  
Location/Qualifiers  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 1174; DB 6; Length 1174;  
Best Local Similarity 100.0%; Pred. No. 3,2e-19;  
Matches 1174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCAGACCCCTCAAGCTGACGATTTCAACCCCGCTCTTTTCTTGCGCCCAAGA 60  
DB 1 GAGCAGACCCCTCAAGCTGACGATTTCAACCCCGCTCTTTTCTTGCGCCCAAGA 60  
QY 61 TCGCTGATCTCCGCTTCTGCGCAACCCGCGCGCTGACCTTCTGCTGACCTCG 120  
DB 61 TCGCTGATCTCCGCTTCTGCGCAACCCGCGCGCTGACCTTCTGCTGACCTCG 120  
QY 121 CGGCGACGAGTGGCAGTGCACGAGATACGAGATCTGCGACCGCGTGG 180  
DB 121 CGGCGACGAGTGGCAGTGCACGAGATACGAGATCTGCGACCGCGTGG 180  
QY 181 GCGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240  
DB 181 GCGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240  
QY 241 TGTGCACTTCAATGTCAGTGGCGTGCACGCGCGCGCTGCGCTACTGCGCG 300  
DB 241 TGTGCACTTCAATGTCAGTGGCGTGCACGCGCGCGCTGCGCTACTGCGCG 300  
QY 301 AGACTCCCTGGCGGCGTGAACGACATCTCCCTCAAGCGCTTGGCGCGACGAGCAT 360  
DB 301 AGACTCCCTGGCGGCGTGAACGACATCTCCCTCAAGCGCTTGGCGCGACGAGCAT 360  
QY 361 GCGCGCGAGATCTGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCG 420  
DB 361 GCGCGCGAGATCTGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCG 420  
QY 421 TCGCGCGAGACGATGCTGACGACGACGACGACGACGACGACGACGACG 480  
DB 421 TCGCGCGAGACGATGCTGACGACGACGACGACGACGACGACGACGACG 480  
QY 481 AGTTGATATGCGCATGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540  
DB 481 AGTTGATATGCGCATGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540  
QY 541 TCGCGCGCTCGCGCGCGCTCAATACGCGCGCATTTGTGCGCGACGATGCT 600  
DB 541 TCGCGCGCTCGCGCGCGCTCAATACGCGCGCATTTGTGCGCGACGATGCT 600  
QY 601 TCGCGCGCGCTCAAGCGCGCGCTGCGCGCGTGGCATCTTCAACGCGCG 660  
DB 601 TCGCGCGCGCTCAAGCGCGCGCTGCGCGCGTGGCATCTTCAACGCGCG 660  
QY 661 ACCGACGTTACGTTGACGACGATGACGCGCGCGCGATGCGTTCGCGCT 720  
DB 661 ACCGACGTTACGTTGACGACGATGACGCGCGCGCGATGCGTTCGCGCT 720  
QY 721 GCAAGCGCAACGACGACCTGCGCGCTTCAACCGCGCGACGCGTGG 780  
DB 721 GCAAGCGCAACGACGACCTGCGCGCTTCAACCGCGCGACGCGTGG 780  
QY 781 GCAAGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840  
DB 781 GCAAGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840  
QY 841 GTGGCTGACGCTGACGATGCGCGCTGACGCGCGCGCGCGCGCGCGCG 900  
DB 841 GTGGCTGACGCTGACGATGCGCGCTGACGCGCGCGCGCGCGCGCGCG 900  
QY 901 CCGTGTCTCTGCGCAACGCTGCGAAGTTGAACGCTACTGCGAGTGCCTTAA 960  
DB 901 CCGTGTCTCTGCGCAACGCTGCGAAGTTGAACGCTACTGCGAGTGCCTTAA 960  
QY 961 CAGCTTTGCGACGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020  
DB 961 CAGCTTTGCGACGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020



QY 222 GATGCCAATCTTCAGAGCGCTGTCCGACTTCAATGTCAGTGGGCTGCAACGGCGGCTCG 281  
 DB 163 AAGCGAAATCTTCAGAGCGCTGTCCGACTTCAATGTCAGTGGGCTGCAACGGCGGCTCG 222  
 QY 282 GCTTACTCTGCGCGCAACGAGTCCCTGGCGGTGAAAGCAATCTGCTTACGGCTTC 341  
 DB 223 GCTTCTCTGCGCGCAACGAGTCCCTGGCGGTGAAAGCAATCTGCTTACGGCTTC 282  
 QY 342 GCGCGACGAGCATCTCGCGCGGTGCTCCGACTTCTGCGGTGCTGCGCTGCTGCGCTC 401  
 DB 283 GCGCGACGAGCATCTCGCGCGGTGCTCCGACTTCTGCGGTGCTGCGCTGCTGCGCTC 342  
 QY 402 ACCCTTACTCTGCGGTGCTCCGCGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTG 461  
 DB 343 ACCCTTACTCTGCGGTGCTCCGCGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTG 402  
 QY 462 GCGCGACTGCGAGTAACTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 521  
 DB 403 GCTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 462  
 QY 522 AAGCGCTGAGCTGCGAGTTCGCGCGGTGCTCCGCGCAAGCAATGCTGCTGCTGCTGCTG 581  
 DB 463 GAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 522  
 QY 582 CCGGACGAGTGCATCTCTTCCCGCGCGGTGCTCCGCGCAAGCAATGCTGCTGCTGCTGCTG 641  
 DB 523 CCGGACGAGTGCATCTCTTCCCGCGCGGTGCTCCGCGCAAGCAATGCTGCTGCTGCTGCTG 582  
 QY 642 TGGTTCAGAAAGCGCGCAACCGGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 701  
 DB 583 TGGTTCAGAAAGCGCGCAACCGGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 642  
 QY 702 ATGCTTCCCGCTGCGGTGCTGCGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 761  
 DB 643 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 702  
 QY 762 CCGGACGAGTGCATCTCTTCCCGCGCGGTGCTCCGCGCAAGCAATGCTGCTGCTGCTGCTG 803  
 DB 703 CCGTTCAGAGCAACGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 762  
 QY 804 GCGCTGCGGTGCGCGCAAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 860  
 DB 763 TCCACCACTGAGCGCGCGCAAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 822  
 QY 861 TGGGCTCAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 920  
 DB 823 TGGGCTCAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 882  
 QY 921 TGCAGAGTGAACGACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
 DB 883 TGCAGAGTGAACGACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 922

RESULT 6  
 A21793 1060 bp RNA linear PAT 20-SEP-1995  
 LOCUS A21793  
 DEFINITION H. insolens mRNA for endonuclease component.  
 ACCESSION A21793  
 VERSION A21793.1 GI:1246872  
 KEYWORDS  
 ORGANISM  
 SOURCE Humicola insolens  
 Humicola insolens  
 Eukaryota; Fungi; Ascomycota; mitospicous Ascomycota; Humicola.  
 REFERENCE  
 1 (bases 1 to 1060)  
 AUTHORS  
 TITLE A CELLULASE PREPARATION COMPRISING AN ENDOGLUCANASE ENZYME  
 JOURNAL Patent: WO 91/17243-A 1 14-NOV-1991;  
 FEATURES  
 1..1060  
 Location/Qualifiers  
 source  
 /organism="Humicola insolens"  
 /mol\_type="unassigned RNA"  
 /strain="DSM 1800"

CDS  
 /db\_xref="taxon:34413"  
 10..927  
 /codon\_start=1  
 /product="endonuclease"  
 /protein\_id="CAA01574.1"  
 /db\_xref="GI:1246872"  
 /db\_xref="EMBL:CAA01574"  
 /translation="MRSSPLPSAVVAALPYLALADGSTRYDCCPSGAKAP  
 VNPVPSGANFORITDPKSGCEBGGVAVSCADQTPMAVNDPDLGPAATIASN  
 ENMCACAEILFTSGVAGKMKVOSTSGDINPDLMIPGAVGIDGCTDPF  
 GELPGORVGISSNRECDREPDALRGCMYRPFKXANPSFSFROYCAELVAT  
 GCRINDENPRANOIPSSSTSSPVNDPSTSTSTSTSSPPVQPTTSSGCTARMAQ  
 CGANSGCTTCVAGSTCTKINDMYHQL"

ORIGIN  
 Query Match 37.3%; Score 437.4; DB 6; Length 1060;  
 Best Local Similarity 70.4%; Pred. No. 5,86-68;  
 Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;  
 55 CCGAGATGCGCTTACTTCCCTTTTTCGCAACCTGCGCGCTGCACTTCTCTGCTG 114  
 DB 5 CCGAGATGCGCTTACTTCCCTTTTTCGCAACCTGCGCGCTGCACTTCTCTGCTG 64  
 QY 115 CCTCCGCGCGCAAGTGGCCAGTCCAGAGTACTGCGGAACTGCTGCAAGCGCTGCT 174  
 DB 65 CCTTGGCCGCTGATGCGAG-----GTCCACCGCTACTGAGACTGCTGCAAGCTTCT 118  
 QY 175 GCGCTTGGCGCGGAAAGCGCGCGCTGAGCCAAACCGCTGCAAGCTGCGCAACTTCC 234  
 DB 119 GCGGCTGGCGCAAGAGGCTCCCGTGAACAGAGCTGCTTTTCTGCAAGCGCAACTTCC 178  
 QY 235 ACCGCTGTCGACTTCAATGCTCAATGCTGCGAGTGA---CGCGGCTGCGCTACTCT 291  
 DB 179 ACCGATATCAAGACTTCAAGCTGCGAGTTCGCGAGTTCGCGAGTTCGCGAGTTCGCT 238  
 QY 292 GCGCGACAGACTCCCTGCGCGGTGAACAGCAATCTGCGCTGCAAGCTTGGCGCGAGCA 351  
 DB 239 GCGCGACAGACTCCCTGCGCGGTGAACAGCAATCTGCGCTGCAAGCTTGGCGCGAGCA 298  
 QY 352 GCATGCGCGCGGTGCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 411  
 DB 299 CTATGCGCGCGCAATGAGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358  
 QY 412 CCGGCTCCGCTGCGCGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 472  
 DB 359 CCGGCTCTGCTGCTGCGAAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 418  
 QY 472 GAGGTAACAGTTCATATGCGCATGCGCGCGCGCGGTGCGGATCTTCAAGCGCTGCA 531  
 DB 419 GCGAGCAACACTTCAATCTCAACATCCCGCGCGCGCGGTGCGGATCTTCAAGCGCTGCA 478  
 QY 532 GCTGCGAGTTCGCGCGCTTCCCGCGCTGCAATGCGCGCGCAATTCGCTGCGCGCACT 591  
 DB 479 CTCCCAAGTTTCGCGCGGTGCGCGCGCGCTGCAAGCGCGCATCTCTCCGCGCAAGCT 538  
 QY 592 GCGATTCCTTCCCGCGCGCTGCAAGCGCGCGCTGCAAGCGCGGTGCTGCTGCTGCTGCA 651  
 DB 539 GCGATTCCTTCCCGCGCGCTGCAAGCGCGCGCTGCAAGCGCGGTGCTGCTGCTGCTGCA 598  
 QY 652 AGCGCGCAACCGCGAGTTCAGTTCGCGAGGTGAGTCCCGCGAGATCTGCTGCTGCTGCT 711  
 DB 599 AGCGCGCAATCCGAGCTTCAAGCTTCCGCTGAGTTCAGTTCGCGAGCTGCTGCTGCTGCT 658  
 QY 712 GCTTGGCTGCAAGCGCGCAAGCACTGCGCTTCCCGCTTCAACCCCGCGAGCGGTG 771  
 DB 659 GCACTCGATGCGCGCGCAAGCAAGCACTTCCCGCTTCAACCCCGCGAGCTGCTGCTGCT 718  
 QY 772 GCAAGG-----TGGACCGCGGCGCGCAAGCTGCACTGCGCGCTGCTGCTGCTGCT 813  
 DB 719 GCACCACTCTCGGCTGCAAGCAAGCTTCAAGCAAGCAAGCAAGCTTCAAGCAAGCTTCA 778  
 QY 814 CCGGCGCAAG---GTCTCCCGCGCGCGAGTGGCTGCAAGCTTCAAGAGTGGGCTGCTGCT 870

Db 779 CGAGCCCGCAGTCCAGCTCCAGCTCCGAGCTGCTGAGAGTGGCTAGT 838  
Qy 871 GCGGTGCGATCGGCTTCAGCGGATGACCACTGTGTCTGACCACTGCGAGAAGT 930  
Db 839 GCGGCGGCAATGGCTGAGCGGCTGACCACTGCTGCTGCGACACTTGGCAGAGA 898  
Qy 931 TGAAGACTACTACTCGAGAGTGCCTTAAC 961  
Db 899 TTATGACTGTATCATCATGACTGCTAGAC 929

RESULT 7  
A23635 1060 bp RNA linear PAT 23-FEB-1995  
LOCUS H. insolens endoglucanase gene.  
ACCESSION A23635  
VERSION A23635.1 GI:832886  
KEYWORDS  
SOURCE Humicola insolens  
ORGANISM Humicola insolens  
REFERENCE 1 (bases 1 to 1060)  
AUTHORS Convents, A.C., Busch, A. and Baeck, A.C.  
TITLE Detergent compositions with high activity cellulase and softening  
clays  
JOURNAL Patent: EP 0495258-A 1 22-JUL-1992;  
THE PROCTER & GAMBLE COMPANY  
FEATURES  
source 1. 1060  
location/Qualifiers  
/organism="Humicola insolens"  
/mol\_type="unassigned RNA"  
/db\_xref="taxon:34413"  
10..927  
/codon\_start=1  
/product="endoglucanase"  
/protein\_id="CAA01681.1"  
/db\_xref="GI:832887"  
/translation="MRSSPLPSAVVAALPYLALADGSTRYMDCKPSGMAKKAP  
VNPVFSNANFORITDFDAKSGCEPBGVAVSCADQTPAVAVNDPFLGPAATSIAGSN  
EAGMCCACVETLFTSGVAKGVVSTSTGSDGSHDLAI PGAGVGFDECTPOF  
GGIPLGRRYGGTSSHNECDREPDLAKGCFRPFKMAANPSFSPRYOCDAELVART  
GCRNDNENPRAVOIPSSSTSSVNPQSTSTSTSTSSPPVQPTPSPGCTAERMAQ  
CGSNWSSGCTTCVAGSTCTKINDMYHGL"

CDS  
sig\_peptide 10..72  
ORIGIN

Query Match 37.3%; Score 437.4; DB 6; Length 1060;  
Best Local Similarity 70.4%; Pred. No. 5.8e-68;  
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

Qy 55 CGAGATGCGCTCTAATCCGCTTCTGACACACCCCTGACCGCTGCACTTCTGCTG 114  
Db 5 CCAAGATGCGCTCTAATCCGCTTCTGACACCCCTGACCGCTGCACTTCTGCTG 64  
Qy 115 CCTCCGCGGCAATGCGCTGCGACAGTCCACAGATCTGAGATCTGCAAGCCGCT 174  
Db 65 CCTTGGCGCTGATGCGAG-----GTGCAACCGGCTACTGAGACTGCTGAGAGCTTGGT 118  
Qy 175 GCGGCTGAGCCCGAGAGGCGCGCTGAGCCCAACCGGTCTAAGCTGCGATGCACTTCC 234  
Db 119 GCGGCTGAGCGCAAGAGGCTCCCTGACACAGCTGTCTTCTTCTGAGAGCACTTCC 178  
Qy 235 AGCGCTGTCCGACTTCAATGTCCAGTGGGCTGCA---CGCGGCTGCGCTTACTCT 291  
Db 179 AGCGTATCGAGACTTCCAGCGCAAGGCTCGGCTGCGAGCCGAGCGGTGCTGCTTGT 238  
Qy 293 GCGCGGACGACCTCCCTGGGCGGTGAGACGAACTGCTAGCTAGCTTGGCTTGGCGGACGA 351  
Db 239 GCGCGGACGACCCCATGGGCTGTGAGACGAACTGCTGCGCTTGGTTGCTGACCACT 298  
Qy 352 GCATGCGCGGCGGTCCGAATCTCTGTGTGCTGCGCTGCTGCAAGGCTCACCTTCACTT 411

Db 299 CTATTGCGGAGAGATGAGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 358  
Qy 412 CGGATCCGCTGCGCGCAAGACATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 471  
Db 359 CGGATCCGCTGCGCGCAAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 418  
Qy 472 GAGTATACGATTTGATATGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 531  
Db 419 GAGGACCACTTGTATCTCAATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 478  
Qy 532 GCTCCGAGTTGGCGCGCTCCCGCGCGCTCAATACCGCGCGCGCGCGCGCGCGCGCG 591  
Db 479 CTCCGAGTTGGCGCGCTCCCGCGCGCTCAATACCGCGCGCGCGCGCGCGCGCGCG 538  
Qy 592 GCGATTTCTTCCCGCGCGCGCTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 651  
Db 539 GCGATGCGTTCCCGCGCGCGCTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 598  
Qy 652 AGCGCGACACCGCGCGCTTCAAGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 711  
Db 599 AGCGCGACACCGCGCGCTTCAAGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 658  
Qy 712 GCTCCG 771  
Db 659 GCGCGGATGCG 718  
Qy 772 GCAAGCG-----TGCGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 813  
Db 719 GCACGCGCTCTCGGTCACACGCGCTTACACGCGCGCGCGCGCGCGCGCGCGCGCG 778  
Qy 814 CGGCGCGAGC---GTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 870  
Db 779 CGAGCCCGCGAGTCCAGCGCTTCAAGCTTCCAGCGCGCGCGCGCGCGCGCGCGCG 838  
Qy 871 GCGGTGCGATCGGCTTCAGCGGATGACACACTGTGTCTGTGACACCGCTGCGAGAAGT 930  
Db 839 GCGGCGGCAATGGCTGAGCGGCTGACACACTGCGCTGCGCGCGCGCGCGCGCGCG 898  
Qy 931 TGAAGACTACTACTCGAGAGTGCCTTAAC 961  
Db 899 TTATGACTGTATCATCATGACTGCTAGAC 929

RESULT 8  
A23644 1060 bp RNA linear PAT 23-FEB-1995  
LOCUS H. insolens endoglucanase gene.  
ACCESSION A23644  
VERSION A23644.1 GI:832890  
KEYWORDS  
SOURCE Humicola insolens  
ORGANISM Humicola insolens  
REFERENCE 1 (bases 1 to 1060)  
AUTHORS Convents, A.C., Busch, A. and Baeck, A.C.  
TITLE Detergent compositions with high activity cellulase and quaternary  
ammonium compounds  
JOURNAL Patent: EP 0495554-A 1 22-JUL-1992;  
THE PROCTER & GAMBLE COMPANY  
FEATURES  
source 1. 1060  
location/Qualifiers  
/organism="Humicola insolens"  
/mol\_type="unassigned RNA"  
/strain="DSM 1800"  
/db\_xref="taxon:34413"  
10..927  
/codon\_start=1  
/product="endoglucanase"  
/protein\_id="CAA01683.1"  
/db\_xref="GI:832891"  
/translation="MRSSPLPSAVVAALPYLALADGSTRYMDCKPSGMAKKAP  
VNPVFSNANFORITDFDAKSGCEPBGVAVSCADQTPAVAVNDPFLGPAATSIAGSN



QY GCTCCGAGTTCCGAGGCGGCTCCCGCGCGCTCAATACGCGGAGATTGGTCGCGGCACAGT 591

Db CTCCCGAGTTCCGAGGCGGCTCCCGCGCGCTCAATACGCGGAGATTGGTCGCGGCACAGT 538

QY GCGATCTCTTCCCGCGCGCTCAAGCCCGGCTGCAGTGGCGGTTGACTGGTTCCAGA 651

Db GCGATCGGTTCCCGCGCGCTCAAGCCCGGCTGCAGTGGCGGTTGACTGGTTCCAGA 598

QY ACGCTCCACAACTCCGACGTTCAAGTTCACAGAGGTGAGTGGCTCCGCGAGTACGTTGCC 711

Db ACGCTCCACAACTCCGACGTTCAAGTTCACAGAGGTGAGTGGCTCCGCGAGTACGTTGCC 658

QY GCTCCGCGGTGCAAGGCGCAAGACGACTCAAGTCTCCCGTTCACATCCCGCCAGCGGTT 771

Db GCACTCCGAGTCCGCGCGCAAGACGAGGCAACTCTCCGCTGACAGATCCCTCCGACGA 718

QY GCAACCG-----TGCACCGGAGACGCCACGCTGACTGCGCTGGGT 813

Db GCACTCAGCTCTCCGGTCAACGAGCTTACAGACACCGACACACCTTCCACTCAACACT 778

QY CGGCGCACAGC---GTCTCCCGCGGCGCGCGAGTGTGACGCTCAGAAAGTGGGCTAGT 870

Db CGAACCCGCAAGCTCAAGCTTCAAGCTCCCGAGCGCTGCACTGCTGAGAGGTGGGCTAGT 838

QY GCGGTGGCATCGGCTTCAAGCGATGCACCAACTGTGTCTTGGCACCACTTCCAGAAAGT 930

Db GCGGCGGCAATGGCTGAGAGCGGCTGCACCACTGCTGCGTGGCAGCACTTCCAGAAAG 898

QY TGAACGACTACTACGACGACATGCTCTTAAAC 961

Db TTATATCTGGTACATCATGATGCTGTATGAC 929

RESULT 10				
A23957	A23957	1060 bp	DNA	linear
LOCUS	Endoglucanase gene.			
DEFINITION	A23957			
ACCESSION	A23957.1			
VERSION	GI:832898			
KEYWORDS				
SOURCE	Humicola insolens			
ORGANISM	Humicola insolens			
REFERENCE	Humicola insolens			
AUTHORS	Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Humicola.			
TITLE	1 (bases 1 to 1060)			
JOURNAL	McCorquodale, F. and Busch, A.			
FEATURES	Dye transfer inhibiting compositions			
source	Patent: EP 0540784-A 1 12-MAY-1993,			
	THE PROCTER & GAMBLE COMPANY			
	Location/Qualifiers			
	1..1060			

```

10..927
/codon_start=1
/product="endoglucanase"
/protein_id="CAA01698.1"
/db_xref="GI:832899"
/db_xref="RENTREMBL:CAA01698"
/tranlsation="MRSSPLPSAVYALPTLALADPSRYWVDCCKDSGGMAKAP
VWQPFSCNNFORITPDPAKSGCEPQGVANASCAQDTPWANVDDPALGRAISLAGNS
EGMGCAGCEBLRTSGPVAQKKNVDSYSTGSDFGNNHDNLITGGGVEIGTGCPPQ
GGIPQRCVGSISRENCDFPDALPKQCGDFWNNHNSFEFRVQCPALVATP
GRRDNDGNPAVOIPSSSTSPVNOPTSTSTSTSTSPVQPTTSPSCTAERMAQ
CGNGMSGCTTCAASCTCTKINDWHQCL"

```

```

Query Match      37.3%;   Score 437.4;   DB 6;   Length 1060;
Best Local Similarity 70.4%;   Pred. No. 5.86-68;
Matches 655;   Conservative 0;   Mismatches 246;   Indels 30;   Gaps 4

```

Db 5 CCAAGATAGCGTCTCTCCCTCCCTCCCTCCCTCCGCTGCGCGTGTGAGCGCGCTGCGAGTATTGG 64

Qy 115 CTTCCGCGCGCATGTGCGAGTGTGCGCATGTCAAGATATCTGGAGCTGTCTGCAAGCGCTGCT 174

Db 65 CCGTTTACCGCTGATGGCAG- - - - -GTCCACCCCGCTACTGGAGCTGTCTGCAAGCGCTTGGT 118

Qy 175 GCGCTTTGGCGCGGAGAGCGCGCGCTGAGCGCAACCGAGTCTACCGGTGCGAGTGGCAACTTCC 234

Db 119 GCGGCTGGGCGCAAGAGCGTCCCTGTAGACCAAGCTGTCTTTTCTGTGCAAGCGCAACTTCC 178

Qy 235 AGGCGCTGTCCGACTTCAATGTCCAGTGGGCGTGCAG- - -CGGCGGCTGGGCGTACTCT 291

Db 179 AGGATATCAACGAGCTTTCAGACGCGCAAGGCTCGGCTGCGAGCGCGGCGGATGTGCCATCTCGT 238

Qy 292 GCGCCGACCAAGATCTCCCTGGGCGGTGAAGCAATCTGCGCTTACCGGCTTGGCGCGACAGA 351

Db 239 GCGCGCACCAAGACCCATAGGATAGGATGTGTGTGAGAGCAACTTGTGCGCTTTGTCTGCAACT 298

Qy 352 GCATCCGCGGCGGAGTCCGATCTCTGTGTGTGCTGCGCTGCTACAGCGCTCACTTCACT 411

Db 229 CTATTGTCCGCGAGATATGAGCGGGCTGTGGTGTGGCTGTCTACAGGCTCACTTCACT 358

Qy 412 CCGGTCCCGTGTCCGCGCAAGACATAGTGTGTGATCAAGAGCATGGGCGGAGACCTGG 471

Db 359 CCGGTCTGTGTGTGCGAAGAGATGTGTGTCACTGTCCACACACTGGGCGGATCTTGG 418

Qy 472 GAAGTACACAGTTCGATATTCGCAATGCGCGCGCGGCGGCGGCGGCACTTTCACACGCGTCA 531

Db 419 GAGAGAACCATCTTGATCTCAAACTCCCGCGCGCGGCGGCGGCACTTTCACAGATGCA 478

Qy 532 GCTCGAGATTTGCGTGGCGCTCCCGCGCTCTAATAGGCGGCACTTTCGTCGCGGAGCACT 591

Db 479 CTCGCCAGTTTGGGCGGCTGTGCCCGCGCGAGCGTCAAGGCGGCACTTGTGCCGCAAGGAT 538

Qy 592 GCGATTTCTTCCCGCGCGCGCTCAAGCGCGCGCTGCAGTGGCGGTTTGACTGTTCAGAG 651

Db 539 GCGATCGATTTCCCGAGAGCGCTCAAGCGCGCGCTGTACTGGCGCTTTCAGTGGTTCAAGA 598

Qy 652 ACGCGACAACTCCGACGTTACGTTCCAGCGAGTCCAGTGCCTCCGCGAGATGCTTGTGCC 711

Db 599 ACGCGACAACTCCAGAGTTCAGCTTCGTCAAGTCCAGTCCAGCGCGAGCTGTGTGCTC 658

Qy 712 GCTCCGCGCGCAAGCGCAACGACATCTCAGAGTTCCTCCGCTTTCACCCGCCCAAGCGGTG 771

Db 659 GCGACCGAGTCCGCGCGCAAGCGACGCGGCACTTCCCTCGTGTCAAGATCCCTCCAGCA 718

Qy 772 GCAACGCG- - - - -TGGACCGGAGCGCCCAAGTGTGATGCTCGCTGGGCT 813

Db 719 GCAACAGGCTTCCGATCAACAGCGCTTACAGAGACCAAGCAACAGTCAAGCTTCAACCT 778

Qy 814 CGGCGCGAGAC- - -GTCTCCCGCGCGGCGCGGAGTGTGACGCTTCAAGAGTGGGCTCAAGT 870

Db 779 GAGACCCGCGAGTCCAGCTTCAAGCTCCCAAGCGGCTGTGACGTGTAGAGGTTGGGCTCAAGT 838

Qy 871 GCGGTGAGCATCGGCTTCAAGCGGATGCAACACTGTGTCTGTGGACACACTGCGCAAGAT 930

Db 839 GCGGCGGCAATAGCTGTAGAGGCGGTGCAACACTGCGGTGCGCGACACTTGCACAGAGA 898

Qy 931 TGAACAGACTACTCTGCGCAAGTGGCTCTTAAC 961

Db 899 TTATATATCGTATCATCATAGTCTGTAGAC 929

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
A41658	Sequence 1 from Patent EP063311.	A41658	A41658.1	GI:2297284	unidentified unidentified unclassified.	unclassified.



REFERENCE 1 (bases 1 to 1060)  
 AUTHORS Harbors, I.M. and Jensen, M.P.  
 TITLE Hydrophobic amines for cellulase stabilization in lipid detergent  
 JOURNAL compositions containing anionic surfactant and cellulase  
 Patent: EP 063311-A 1 11-JAN-1995;  
 PROCTER & GAMBLE (US)  
 COMMENT Other publication CA 2165771 950105.  
 FEATURES  
 Location/Qualifiers  
 source 1..1060  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"

ORIGIN  
 Query Match 37.3%; Score 437.4; DB 6; Length 1060;  
 Best Local Similarity 70.4%; Pred. No. 5.8e-68;  
 Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

Db 55 CCAAGATGCTCTACCTCCGTTCTTGGACAAACCTGGCCGCTGCACTTCTCTGTGG 114  
 5 CCAAGATGCTCTACCTCCGTTCTTGGACAAACCTGGCCGCTGCACTTCTCTGTGG 64  
 Qy 115 CCTCCGCGCAGTGGCAGTGGCCAGTGGCAAGATATGAGAGTGGAGTGGCAAGCCGCT 174  
 Db 65 CCTTGGCGCTGATGGCAG-----GTCCACCCGCTACTGGGACTGGCAAGGCTTGGT 118  
 Qy 175 GCGCTTGGCGCGGAGAGCGCGCGTGGACCAACGCTGTAACGCGTGGCAATCTTC 234  
 Db 119 GCGGCTGAGGCAAGAGGCTCCGCTGAACCAAGCTGTCTTCTTCCGCAACCCCACTTCC 178  
 Qy 235 AGCGGCTGTCCGACTTAATGTTCAGTGGGCTGGCA---CGCGCGCTCGGCGCTACTCCT 291  
 Db 179 AGGATTCACGACCTTGAACGCGAGTCCGCTGGCGAGCGCGGCTGGCTACTCCT 238  
 Qy 292 GCGCGACCAAGACTCCCTGGCGGAGTGAACCAATCTCCGCTACGCGCTTGGCGAGCA 351  
 Db 239 GCGCGACCAAGACTCCCTGGCGGAGTGAACCAATCTCCGCTACGCGCTTGGCGAGCA 298  
 Qy 352 GCAATGCGCGGCGGAGTCCGATCTCGTGGTGGTGGCTGGCTTACGCGCTCACTTCACTT 411  
 Db 299 CTATTCGCGGAGCAATGAGGCGGCTGGTGGTGGCTGGCTTACGCGCTCACTTCACTT 358  
 Qy 412 CGGTCGCGTGGCGGAGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 471  
 Db 359 CGGTCGCTGTGGAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 418  
 Qy 472 GAAGTACCAAGTTCGATATGCGCATGCCGCGCGCGCTGGGCACTTTCACACGCGCTGCA 531  
 Db 419 GAGCAACCACTTCGATTCATATCCCGGCGCGCGCTGGGCACTTTCGAGAGTGA 478  
 Qy 532 GCTGCACTTGGCGGCTTCCCGCGCGCTCAATGCGCGCACTTTCGTCGCGAGCACT 591  
 Db 479 CTCCTCCAGTTCCGCGGCTTCCCGCGCGCTCAATGCGCGCACTTTCGTCGCGAGCACT 538  
 Qy 592 GCGATTCCTTCCCGCGCGCTCAAGCCGCGCTGCGAGTGGCGGTTTACCTGGTTCAGA 651  
 Db 539 GCGATGCTTCCCGAGCGCTCAAGCCGCGCTGCTAATGCGCTTTCGAGTTCAGAA 598  
 Qy 652 AGCGCGACCAACCGACCTTCAAGCTTCAAGAGTGGAGTGGCGCGCGAGATGATTCGCC 711  
 Db 599 AGCGCGACCAACCGACCTTCAAGCTTCCGAGGTCAGAGTGGCGCGAGCTGTCGCTC 658  
 Qy 712 GCTTGGCGCTCAAGCGCAAGAGTCAAGTTCCTCCCTTTCACCCCGTCAAGGCGTG 771  
 Db 659 GACCGGATGCGCGCGCAAGAGGCAATCTTCCCTGCTGCTCAAGATTCCTTCAGGA 718  
 Qy 772 GCAAGG-----TGGACAACCGGACCGCCACGCTGCACTGCGCTGGGT 813  
 Db 719 GACCAAGCTTCGCGTCAACAGCTTACAGACCAAGACCAAGTTCACCTTCACCACT 778  
 Qy 814 GGGCGCAAG---GTCTCCGCGCGCGCAAGTGGCTGCAAGTCTCAAGAGTGGGCTCACT 870  
 Db 779 GAGCGCGCAAGTCCAGCTTACGACTCCAGCGGCTGCACTGCTGAGAGTGGGCTCACT 838

Qy 871 GCGGTGGCAATGCGCTTACGCGGATGACCACTGTGTCTTGGACCACTGCGAGAGT 930  
 Db 839 GCGCGGCAATGCGCTGAGCGCGTGGACCACTGCGTGGCTGGGCACTTGCAGAGGA 898  
 Qy 931 TGAACGACTACTGCTGAGTGGCTTCAAC 961  
 Db 839 TTAATGACTGTTCATCACTAGTCCCTGTAGAC 929

RESULT 12  
 A68060 1060 bp DNA linear PAT 05-MAY-1999  
 A68060  
 LOCUS Sequence 55 from Patent W09743409.  
 DEFINITION A68060  
 ACCESSION A68060.1 GI:4756862  
 VERSION  
 KEYWORDS  
 SOURCE Humicola insolens  
 ORGANISM Humicola insolens  
 Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Humicola.  
 REFERENCE 1 (bases 1 to 1060)  
 AUTHORS Dalboege, H., Diderichsen, B., Sandal, T. and Kauppinen, S.  
 TITLE METHOD OF PROVIDING NOVEL DNA SEQUENCES  
 JOURNAL Patent: WO 9743409-A 55 20-NOV-1997;  
 NOVONORDISK AS (DK)  
 FEATURES  
 Location/Qualifiers  
 source 1..1060  
 /organism="Humicola insolens"  
 /mol\_type="unassigned DNA"  
 /strain="DSM 1800"  
 /db\_xref="taxon:34413"  
 10..927  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CA842307.1"  
 /db\_xref="GI:4756863"  
 /translation="MRSSPLPSAVVAALPVALAANDGRSTRVWCCPSGMAKKAP  
 VNPVPSGANPFRITTPDKSGCTGCGYANSCADDPYAVNDPLSPATISASGN  
 EAMGCAACEYELFTSGPVAKRNVOSTSGDLSNHPUNTPGGGVTGDCETPE  
 GGLPGRYGELISRNEDCRFPDLAKGCMFDFWPNXADPSPFRQVCPAEIVAT  
 GGRNDGNFPAVOIPSSSTSPVNPOTSTSTSTSTSTSTSPVQPTTSGCTARMAQ  
 CGNGSGTTCTVAGSTCTKINDWHQCL"  
 10..72  
 sig\_peptide  
 mat\_peptide  
 73..924  
 /product="unnamed"

ORIGIN  
 Query Match 37.3%; Score 437.4; DB 6; Length 1060;  
 Best Local Similarity 70.4%; Pred. No. 5.8e-68;  
 Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

Qy 55 CCAAGATGCTCTACCTCCGTTCTTGGACAAACCTGGCCGCTGCACTTCTCTGTGG 114  
 Db 5 CCAAGATGCTCTACCTCCGTTCTTGGACAAACCTGGCCGCTGCACTTCTCTGTGG 64  
 Qy 115 CCTCCGCGCAGTGGCAGTGGCCAGTGGCAAGATATGAGAGTGGAGTGGCAAGCCGCT 174  
 Db 65 CCTTGGCGCTGATGGCAG-----GTCCACCCGCTACTGGGACTGGCAAGGCTTGGT 118  
 Qy 175 GCGCTTGGCGCGGAGAGCGCGCGTGGACCAACGCTGTAACGCGTGGCAATCTTC 234  
 Db 119 GCGGCTGAGGCAAGAGGCTCCGCTGAACCAAGCTGTCTTCTTCCGCAACCCCACTTCC 178  
 Qy 235 AGCGGCTGTCCGACTTAATGTTCAGTGGGCTGGCA---CGCGCGCTCGGCGCTACTCCT 291  
 Db 179 AGGATTCACGACCTTGAACGCGAGTCCGCTGGCGAGCGCGGCTGGCTACTCCT 238  
 Qy 292 GCGCGACCAAGACTCCCTGGCGGAGTGAACCAATCTCCGCTACGCGCTTGGCGAGCA 351  
 Db 239 GCGCGACCAAGACTCCCTGGCGGAGTGAACCAATCTCCGCTACGCGCTTGGCGAGCA 298  
 Qy 352 GCAATGCGCGGCGGAGTCCGATCTCGTGGTGGTGGCTGGCTTACGCGCTCACTTCACTT 411



**TITLE** Packar, S. Maant., Hagen, F., Hjort, C. Malland, and Hastrup, S.  
**JOURNAL** Cellulase preparation comprising an endoglucanase enzyme  
**FEATURES** Patent: US 5946672-A 1 07-SEP-1999;  
Location/Qualifiers

ORIGIN

Query Match	37.3%;	Score 437.4;	DB 6;	Length 1060;
Best Local Similarity	70.4%;	Pred. No. 5.8e-68;		
Matches 655; Conservative	0;	Mismatches 246;	Indels 30;	Gaps 4

QY	55	CCAGAGATGAGCTTAACTCCCGTCTTTCGCAAAACCCGCGCGCTCAACTTCCTCTGAGCG	11.4
Db	5	CCAGAGATGAGCTTCCCTCCCTCTCTCCGTCGCGCTGTGCGCGCCCTCCCGGATTTGG	64
QY	115	CTTCGCGGACAGATGGGAGTGGACAGCTCGATCGAACAGATATCTGGGACTGCTGCMAAGCTGT	174
Db	65	CCCTTGCCTCTAATATGGAG-----GTCAACCGGCTACTAGGGAGCTGGCTGGAACTTGTGT	118
QY	175	GGCGTTGAGCCGGGAAAGCGCGCGTACGACGAACCGGTACGACGCGTGGAGATGCAACTTTC	234
Db	119	GGCGCTGGGACCAAGAAAGCTCCGTTGAACAGACTGTCTTTTCTTGAAAGCAACTTTC	178
QY	235	AGGCGCTGTCCGACTTCAATATGTCAGATGAGGCTGCA---CGGCGGCTGGCGTACTCT	291
Db	179	AGGGTATCAAGGACTTTCGACGCGCAAGTCCGGCTCCGACGCGGAGCGATGTGACTACTGT	238
QY	292	GGCGGCAACAGATCTCCGCGGCGGTGAACAGCAATCTGGCTAGAGCGTCTGGCGGAGAGA	351
Db	239	GGCGGCAACAGACCCCAATGGGCTGTGAACAGACGACTTGGCGGTGTGTCGCACTT	298
QY	352	GCAATCCGCGGCGAGGTCGCAATCTGTGATCTGACGCTGTGAAGCGCTCACTTCACTT	411
Db	299	CTATATGCGCGGAGCAATGAGGCGGCGCTGTGTCTGCGCTGTCAAGACTCACTTCAAT	358
QY	412	CCGAGTCCGCTGCGCGGAGACATATGTGTGTGTCAGTCAAGAGCATGAGCGGCACTTG	471
Db	359	CCGAGTCTGTGTGCTGAGAAAGATATGTGTGTGTCAGTCAAGAGCATGAGCGGCTGTG	418
QY	472	GAAATACCAAGTGGATATATGACATGCGCGCGCGCGCGGTGGAGCATCTTCAAGCATCA	531
Db	419	GCGCAACCACTTGAATCTTCAATCTCCCGGCGCGCGGTGCGCAATCTTCAAGAGATCA	478
QY	532	GCTTCGAGTTCGCGGCGCTCCCGGCGCTCAATACGAGCGCAATTTGGTGGCGGACAGT	591
Db	479	CTCCCGAGTTGGCGGCTGTGCTCCCGGCGAGCGCTTACGCGGCAATCTGGTCCGACAGAT	538
QY	592	GCGATTCCTTCCCGCGCGCTTCAGCCCGGCTGCGAGTGGGTTGATGTGTTCCAGA	651
Db	539	GCGATGAGTTCCCGGAGGCTTCAGCCCGGCTGTCTATGAGCGCTGTGACCTGATTCAGAA	598
QY	652	ACGCGCAACACCGACCTTTCAGCTTTCAGAGGTGAGTGGCCCGCGAGATGTTGCGC	711
Db	599	ACGCGCAACATCCAGAGCTTCAAGCTTCCGTGAGTTCAGTGGCCCAAGCACTGTGTGTC	658
QY	712	GCTTCGAGCTGCAAGCGCAAGACATCCAGCTTCCCGCTTCAACCCCGCAAGCGTGT	771
Db	659	GCAACGAGATGCGCGGCAAGAGAGGCACTTCCCTGCGCTTCAGATCCCTTCACAGA	718
QY	772	GCAACG-----TGGGACCGGAGGCGCAAGTCAATGCGCTGGGT	813
Db	719	GCAACAGCTCTCCGATCAACCAAGCTTACAGAGTACAGACACAGTCCACCTTCCACACT	778
QY	814	CGGCGCAGAC---GTCTTCGAGCGGCGGAGTGGCTGCAAGTCTCAAGTGGGCTCAAT	870
Db	779	CGAGCCCGCAATCCAGGCTACGACTCCCAAGCGCTCACTGTGAGGAGTGGGCTCAAT	838
QY	871	GGGATGGCATTCGCTTCAAGAGGATATCAACCACTGTGTCTTGGACCACTTGCAGAAAT	933
Db	839	GGGCGCGCAATGGCTGAGAGGAGCTGACCACTCTGTGTGAGGAGCACTTGTGACAGAA	898

```

QY      931 TGAACGACTACTCCAGTGCCTTAAC 961
          ||| ||| ||| ||| ||| ||| |||
Db      899 TTATGACTGTACCACTAGTGCTGTAGAC 929

```

RESULT 15

LOCUS	AR163160	1060 bp	DNA	linear	PAT 17-OCT-2001
DEFINITION	Sequence 55 from patent US 6270968.				
REFERENCE	Sequence 55				

ACCESSION	ARI63160
VERSION	ARI63160.1
	GI:16233662

KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.

ORGANISM UNKNOWN.  
Unclassified.  
REFERENCE 1 (pages 1 to 1060)

**AUTHORS**  
Diderichsen, B. (1908)  
Dalb. O. (1908)  
Blashed, H. (1908)

TITLE	Method of providing a h interest
-------	-------------------------------------

JOURNAL	Patent: US 6270968-A 55
FEATURES	Location/Quali

```
source      1. .1060
/organism="unk
```

```
ORIGIN
      /mol_type="unassigned DNA"
```

## ORIGIN

Query Match	37.3%;	Score 437.4;	DB 6;	Length 1060;
Best Local Similarity	70.4%;	Pred. No. 5.8e-68;		
Matches 655;	Conservative 0;	Mismatches 246;	Indels 30;	Gaps 4;

Qy	55	CGAGATGCGCTACATCCCGGTTCTTGCGAACACCTGGCCGCTGCACATCTCCCTGAGTGG	114
Db	5	CGAAGATGCGTTCTCCGCCCTCTCTCCGTCGCGCTTGTGTGGCCGCCCTCCGGGTGG	64
Qy	115	CTTCGCGCGCGAGTGGCGATGGCGAGTGCACGAGATACCTGTACTGCTGTAAGCCGTGT	174
Db	65	CCCTTCGCCCTGATGGCAG-----GTCCACCGGCTACCTGCGGATCTGTGCAAGCCTTGT	118
Qy	175	GCGCTGGCGCGGGAGAGCGCGCGCTCAGCGCACCGGCTCTAGCGCTGTGATGCAACTTCC	234
Db	119	GCGGCTGGCGCGAGAGAGGCTCCCGTGAACAGGCTGTCTTTTCTTGCAAGCGCACTTCC	178
Qy	235	AGCGCCTGTCGACTTCAATGTTCAGTGGGCTGCA--CGGCGGCTCGGCTTACTCT	291
Db	179	AGGTTATCAGGACTTGCACGCGCAAGTCCGCTCGGCGAGCGCGGCGGTGTGCTCTGT	238
Qy	292	GCGCGCACCAATCCTCCGCGGGGCTGAAACCAAAATTCGCTCAAGGCTCTGCGCGAGCA	351
Db	239	GCGCGCACCAAGCCCAATGGGCTGTGAACACACACTTGGCGCTGAGTTTGTCGACACT	298
Qy	352	GCATTCGCGCGGAGTCCGATCTCTGTGGGTGCTGCGCTGCTACGCGCTACCTTCACTT	411
Db	299	CTATTTCGCGGACGAATGAGAGCGGAGCTGTGTGCTGCGCTGCTACGACTTCACTTCAAT	358
Qy	412	CCGCTCCCGCTCGCGGGAAGAAACAATGGTGTGAGTGCACAGAGCACTGGGCGACCGG	471
Db	359	CCGCTCCCTGTGTGCGGGAAGAGATGTCTGCATGTGCACAGCACTGGGCGAGTCTTG	418
Qy	472	GAAATACCAAGTTGATATGCGCATGCCCGGCGCGGAGTGGGATCTTCAACGCGCTGCA	531
Db	419	GCAGCAACCACTTGATCTCAACATCTCCCGCGGCGGCTCGGCACTTTCGACGAGATCA	478
Qy	532	GCTGCGAATGGGGGCGCTCCCGCGCGCTCAATACGCGGCGCATTTGTGTGCGGACAGT	591
Db	479	CTCCCAAGTTGGGCGTGTCCCGGCGAGGCTAGGCGGCGCATCTGTGTCCCGAACAGT	538
Qy	592	GCGATTCCTTCCCGCGCGCTCAAGCCCGGCTGCGCAGTGGCGGTGTTGACTGTATTCAGA	651
Db	539	GCGATGGGTTCGCCBACGCCCTTCAGCCCGGCTGTACTGGGCTTGCATGGTATCAAGA	598
Qy	652	ACGCGCAGACCCGACCTTCACTTCCAGCGAGGTGAGTGGCCCGCGAGATCTGTTGCC	711

```
Db      539  AGGCCGACAAATCCGAGCTTCAAGCTTCGGTCAGGTCCAGTGCACGCGAGCTCTGGCTC 658
Qy      712  GCTCGGGCTGCAGGCGCAACGACATCCAGCTTCCCGCTTTCACCCCCCAAGCGGTG 771
Db      659  GCACCGAATGCCCGGCACACACACGCGCACTTCCCTGCGGTCCAGATCCCTTCAGCA 718
Qy      772  GCMAAGG-----TGGACCGGGAAGCCCAAGTCGACTGCGCCTGAGT 813
Db      719  GCACCAAGCTCTCCGCTCAACCAAGCTTACAGCACACAGCACACGTCACCTCCACCACT 778
Qy      814  CGGGCCAGAC--GCTCTCCGCGCGCGGCGAGTGGCTGCACGCTCAGAAATGGGCTCAGT 870
Db      779  CGAGCCGCGCAGTCCAGCTTACGACTCCCGAGCGGCTGCACTGCTGAAAGGTGGGCTCAGT 838
Qy      871  GCGGTGGCATCGGCTTCAAGCGGATGCACACCTGTGTCTTGGACACCCTGCAGAGT 930
Db      839  GCGCGGCAATGGCTGAGCTGGTGCACCACTGCGTGGCTGGGAGCACTTGCAGAGAA 898
Qy      931  TGAAGACTACTACTCGAGTGGCTTAAAC 961
Db      899  TTAATGACTGGTACCATCAGTGCCTGTAGAC 929
```

Search completed: July 7, 2004, 10:19:30  
Job time : 7435 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using *sw* mode:

Run on: July 7, 2004, 08:03:07 ; Search time 716 Seconds  
(without alignments)  
6965.621 Million cell updates/sec

Title: US-10-007-521-11

Perfect score: 1174  
Sequence: 1 GAGGAGCAGCCCTCAAGCTG.....AAAAAAAAAAAAAAAAAAAA 1174

Scoring table:  
IDENTITY NTC  
Gapop 10.0, Gapext 1.0

Searched: 3373863 segs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seg length: 0  
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: Geneseg 29Jan04:\*  
2: geneseg1980s:\*  
3: geneseg1990s:\*  
4: geneseg2000s:\*  
5: geneseg2001as:\*  
6: geneseg2001bs:\*  
7: geneseg2002as:\*  
8: geneseg2003bs:\*  
9: geneseg2003cs:\*  
10: geneseg2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	* Match Length	ID	Description
1	1174	100.0	1174 2 AAT39050	Aat39050 cDNA enco
2	1174	100.0	1174 2 AAV39050	Aav39050 Monocompo
3	453.6	38.6	1261 2 AAV23748	Aav23748 Humicola
4	451.6	38.5	922 2 AAV15072	Aav15072 Hybrid DN
5	439	37.4	1060 2 AAZ60178	Aaz60178 H. insol
6	437.4	37.3	1058 2 AAQ26405	Aaq26405 Humicola
7	437.4	37.3	1060 2 AAQ14856	Aaq14856 Humicola
8	437.4	37.3	1060 2 AAQ25932	Aaq25932 Cellulase
9	437.4	37.3	1060 2 AAQ30067	Aaq30067 Sequence
10	437.4	37.3	1060 2 AAQ29934	Aaq29934 Endogluc
11	437.4	37.3	1060 2 AAQ49941	Aaq49941 Endogluc
12	437.4	37.3	1060 2 AAQ41732	Aaq41732 Dye trans
13	437.4	37.3	1060 2 AAT10182	Aat10182 Alkaline
14	437.4	37.3	1060 2 AAV15065	Aav15065 Humicola
15	437.4	37.3	1060 2 AAV16102	Aav16102 Humicola
16	437.4	37.3	1060 2 AAAS3334	AAas3334 Endogluc
17	438.8	37.1	1060 2 AAQ26380	Aaq26380 Endogluc
18	431	36.7	1060 2 AAQ30072	Aaq30072 43RD endo
19	429.4	36.6	1060 2 AAQ31181	Aaq31181 H. insol
20	412.8	35.2	922 2 AAV15073	Aav15073 Hybrid DN
21	393.4	33.5	925 2 AAV15076	Aav15076 Hybrid DN
22	390.6	33.3	807 2 AAV16104	Aav16104 Humicola
23	377.2	32.1	1257 2 AAV19378	Aav19378 Humicola

24	377.2	32.1	1257 2 AAV13840	Aav13840 Humicola
25	377.2	32.1	1257 6 AAL43251	Aal43251 Humicola
26	377.2	32.1	1257 6 ABL60700	AbL60700 H. insol
27	377.2	32.1	1257 6 ABL59235	AbL59235 Nucleoid
28	377.2	32.1	1257 8 ADA37115	Ada37115 Humicola
29	368.2	31.4	1154 2 AAT39048	Aat39048 cDNA enco
30	366.4	31.2	928 2 AAV15074	Aav15074 Hybrid DN
31	365.6	31.1	927 2 AAV29596	Aav29596 Humicola
32	353.6	30.1	936 2 AAT66541	Aat66541 20K-cellu
33	337.6	28.8	915 2 AAV15075	Aav15075 Hybrid DN
34	337	28.7	894 2 AAT39061	Aat39061 Chimeric
35	334.8	28.5	927 2 AAT39062	Aat39062 Fusarium
36	315.2	26.8	1473 2 AAQ14857	Aaq14857 Fusarium
37	315.2	26.8	1473 2 AAQ25933	Aaq25933 Cellulase
38	315.2	26.8	1473 2 AAQ26407	Aaq26407 Fusarium
39	315.2	26.8	1473 2 AAQ26382	Aaq26382 Endogluc
40	315.2	26.8	1473 2 AAQ29935	Aaq29935 Endogluc
41	315.2	26.8	1473 2 AAQ49942	Aaq49942 Endogluc
42	315.2	26.8	1473 2 AAQ60179	Aaq60179 F. oxyspo
43	315.2	26.8	1473 2 AAV16103	Aav16103 Fusarium
44	313.6	26.7	885 2 AAT39075	Aat39075 Chimeric
45	313.6	26.7	1473 2 AAQ41733	Aaq41733 Dye trans

#### ALIGNMENTS

RESULT 1	
AAAT39050	
ID AAT39050 standard; cDNA; 1174 BP.	
XX	
AC AAT39050:	
DT 20-MAY-1997 (first entry)	
XX	
DE cDNA encoding cellulytic enzyme #4 of the invention.	
XX	
KM Cellulytic enzyme; endoglucanase; hydrolisis; cellulose; microorganism;	
KM plant cellulase; catalytic region; textile; backstaining; bio-polishing;	
KM stone-washing; cellulosic fabric; colour clarification; defibration;	
KM cell wall degradation; paper pulp; debarking; fibre modification;	
KM enzymatic de-inking; drainage improvement; ss.	
XX	
OS Thielavia terrestris.	
XX	
FH Key	Location/Qualifiers
FT CDS	59..959
FT	/*tag= a
FT	/product= "endoglucanase"
XX	
W09629397-AL.	
XX	
PD 26-SEP-1996.	
XX	
PF 18-MAR-1996; 96WO-DK000105.	
XX	
PR 17-MAR-1995; 95DK-00000272.	
PR 08-AUG-1995; 95DK-00000885.	
PR 08-AUG-1995; 95DK-00000886.	
PR 08-AUG-1995; 95DK-00000887.	
PR 08-AUG-1995; 95DK-00000888.	
PR 12-FEB-1996; 96DK-00000137.	
XX	
PA (NOVO ) NOVO-NORDISK AS.	
XX	
PI Schuelein M, Andersen LN, Laessen SF, Kauppinen MS, Lange L;	
XX Nielsen RI, Ihara M, Takagi S;	
DR WPI: 1996-443173/44.	
XX P-PSDB; AAM04928.	
XX	
PT New endo-glucanase enzyme preparations - confg. conserved catalytic	
PT regions, useful for treating fabrics, textiles, plant material or paper	

pulp.  
 Claim 72; Page 160-161; 316pp; English.  
 AAT39047-139054 represent the coding sequences for the enzymes of the invention. The encoded enzymes possess cellulolytic (particularly endoglucanase) activity. Cellulolytic enzymes are involved in the hydrolysis of cellulose, and are synthesised by a large number of microorganisms and plants. The enzymes of the invention containing the conserved catalytic regions (such as AAM04913) exhibit improved performance, e.g. 50 times higher performance, compared to multiple domain enzymes. The enzymes can be used for the treatment of fabrics or textiles, preferably for preventing backstaining, for bio-polishing or for stone-washing cellulosic fabric. They can also be used to provide colour clarification for laundry. The enzymes can also be used for the degradation or modification of plant material, such as cell walls. They can also be used in the treatment of paper pulp preferably for debarking, defibration, fibre modification, enzymatic de-inking or drainage improvement.  
 Sequence 1174 BP; 243 A; 395 C; 320 G; 216 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1174; DB 2; Length 1174;  
 Best Local Similarity 100.0%; Pred. No. 9e-204;  
 Matches 1174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 GAGGAGCAACCCCTCAAGCTGTAAGTTTCAACCCGCTCTCTTTCTGCGCCCAAGAA 60  
 1 GAGGAGCAACCCCTCAAGCTGTAAGTTTCAACCCGCTCTCTTTCTGCGCCCAAGAA 60  
 61 TGGGCTCTACTCCCGCTTCTTGCAACCCGCGCTGCACTTCCCTGCGCTGCTCG 120  
 61 TGGGCTCTACTCCCGCTTCTTGCAACCCGCGCTGCACTTCCCTGCGCTGCTCG 120  
 62 TGGGCTCTACTCCCGCTTCTTGCAACCCGCGCTGCACTTCCCTGCGCTGCTCG 120  
 121 CGGCGAGTGGAGTGGCAAGTCAAGATATCTGAGATCTGAGACCGCTGCTGCT 180  
 121 CGGCGAGTGGAGTGGCAAGTCAAGATATCTGAGATCTGAGACCGCTGCTGCT 180  
 122 CGGCGAGTGGAGTGGCAAGTCAAGATATCTGAGATCTGAGACCGCTGCTGCT 180  
 181 GGGCGGGAAGGCGCGCTGAGCAACCGCTGCAACCGCTGCAACCGCTGCAACCG 240  
 181 GGGCGGGAAGGCGCGCTGAGCAACCGCTGCAACCGCTGCAACCGCTGCAACCG 240  
 241 TGTCCGACTTCAATGTCAGTGGCTGCAACCGCGCTGCAACCGCTGCAACCGCT 300  
 241 TGTCCGACTTCAATGTCAGTGGCTGCAACCGCGCTGCAACCGCTGCAACCGCT 300  
 242 TGTCCGACTTCAATGTCAGTGGCTGCAACCGCGCTGCAACCGCTGCAACCGCT 300  
 301 AGACTCCCTGGGCGTGAAGCAATATCTGCTTCAACCGCTGCAACCGCTGCAAC 360  
 301 AGACTCCCTGGGCGTGAAGCAATATCTGCTTCAACCGCTGCAACCGCTGCAAC 360  
 302 AGACTCCCTGGGCGTGAAGCAATATCTGCTTCAACCGCTGCAACCGCTGCAAC 360  
 361 GGGGCTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 361 GGGGCTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 362 GGGGCTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 421 TGGCGGCAAGCAATATGTCAGTGGCTGCAACCGCGCTGCAACCGCTGCAACCG 480  
 421 TGGCGGCAAGCAATATGTCAGTGGCTGCAACCGCGCTGCAACCGCTGCAACCG 480  
 422 TGGCGGCAAGCAATATGTCAGTGGCTGCAACCGCGCTGCAACCGCTGCAACCG 480  
 481 AGTTGATATGTCAGTGGCTGCAACCGCGCTGCAACCGCTGCAACCGCTGCAAC 540  
 481 AGTTGATATGTCAGTGGCTGCAACCGCGCTGCAACCGCTGCAACCGCTGCAAC 540  
 541 TGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 541 TGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 542 TGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 601 TGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 601 TGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 602 TGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 661 ACCCGAGTTTCAAGTTCAGAGTGGCTGCAACCGCGCTGCAACCGCTGCAACCG 720  
 661 ACCCGAGTTTCAAGTTCAGAGTGGCTGCAACCGCGCTGCAACCGCTGCAACCG 720

QY 721 GCAAGCCGCAAGAGTTCAGCTTCCCGCTTCAACCCCGCAAGCGTGGCAAGG 780  
 DB 721 GCAAGCCGCAAGAGTTCAGCTTCCCGCTTCAACCCCGCAAGCGTGGCAAGG 780  
 QY 781 GCAAGCCGCAAGAGTTCAGCTTCCCGCTTCAACCCCGCAAGCGTGGCAAGG 840  
 DB 781 GCAAGCCGCAAGAGTTCAGCTTCCCGCTTCAACCCCGCAAGCGTGGCAAGG 840  
 QY 841 GTGGTGAACGCTCAGAGTGGCTGCAACCGCGCTGCAACCGCTGCAACCGCT 900  
 DB 841 GTGGTGAACGCTCAGAGTGGCTGCAACCGCGCTGCAACCGCTGCAACCGCT 900  
 QY 901 CTTGTCTCTGCGCAACCTGCGCAAGTGAACCTTCAACCGCTGCAACCGCTGCA 960  
 DB 901 CTTGTCTCTGCGCAACCTGCGCAAGTGAACCTTCAACCGCTGCAACCGCTGCA 960  
 QY 961 CACTTTTGGCAAGGTTGGCGGCAACCGCAAGAGTGAACCTTCAACCGCTGCA 1020  
 DB 961 CACTTTTGGCAAGGTTGGCGGCAACCGCAAGAGTGAACCTTCAACCGCTGCA 1020  
 QY 1021 TTTTGAAGGCTCAATACATACATACATACATACATACATACATACATACAT 1080  
 DB 1021 TTTTGAAGGCTCAATACATACATACATACATACATACATACATACATACAT 1080  
 QY 1081 TCTCAACCGCACTTGGCGGCGCAATACAGGCGCTTTTCAACCGCTGCAAC 1140  
 DB 1081 TCTCAACCGCACTTGGCGGCGCAATACAGGCGCTTTTCAACCGCTGCAAC 1140  
 QY 1141 AAA 1174  
 DB 1141 AAA 1174  
 RESULT 2  
 AAV39096  
 ID AAV39096 standard; cDNA; 1174 BP.  
 AC AAV39096;  
 DT 21-SEP-1998 (first entry)  
 XX  
 DE Monocomponent endoglucanase encoding cDNA.  
 XX  
 KM Monocomponent endoglucanase; cellulolytic enzyme; garment; abrasion;  
 KM abraded looking jeans; fungus; Thielavia terrestris; stone-washed; ss.  
 XX  
 OS Thielavia terrestris.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 60..959  
 FT /tag=a  
 FT /product="monocomponent endoglucanase"  
 PN EP843041-A1.  
 PD 20-MAY-1998.  
 XX  
 PF 30-MAY-1997; 97BP-00610021.  
 XX  
 PR 13-NOV-1996; 96DK-00001276.  
 XX  
 PA (NOVO) NOVO-NORDISK AS.  
 XX  
 PI Lund H, Kalum L;  
 XX  
 DR WPI; 1998-263308/24.  
 XX  
 PT P-PSDB; AAM63624.  
 XX  
 PS Pair of jeans - with abraded look.  
 XX  
 PS Claim 17; Page 9-10; 16pp; English.



CC residue. The present sequence represents the cDNA sequence of Humicola  
 CC grisea cellulase from the present invention. The new cellulase enzymes  
 CC can be used in e.g. detergent or fabric softener compositions, for bio-  
 CC polishing of new fabrics, for promoting a stone-washed look to cellulose  
 CC containing fabric, for pulp and paper applications, e.g. for debarking,  
 CC defibrillation, fibre modification, drainage improvement, inter fibre  
 CC bonding or for degradation of plant material e.g. for improving feed  
 CC value. The cellulase variants have improved properties with respect to  
 CC e.g. catalytic activity, altered sensitivity to anionic tensides, pH  
 CC optimum or activity profile or stability

XX Sequence 1261 BP; 246 A; 417 C; 336 G; 262 T; 0 U; 0 Other;

Query Match 38.6%; Score 453.6; DB 2; Length 1261;

Best Local Similarity 74.9%; Pred. No. 3.8e-73;

Matches 583; Conservative 0; Mismatches 189; Indels 6; Gaps 1;

```

QY 26 TTTCACCCCGCTCTCTTTCTTGGGCCCCGAGAGTGGGCTTAACTCCGCTCTTGGAC 85
DB 18 TTTTATTCAGCTCATTTATTTTAAACTTCAATATGCGCTCTGCTCAATTTTCGAC 77
QY 86 AACCTGCGCGCTGCACTTCTCTGCTGCTCCGCGGCAATGGAGTGGCACTGCAC 145
DB 78 GGCCCTGCGCGCTGCGCTCCCTTTCGCGAGTCCGCGCGGATGGCA-----ACTGAC 131
QY 146 GAGATTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCGGTGACCA 225
DB 132 CAGATTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 191
QY 206 ACCGATTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 265
DB 192 GCTTGTCTTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 251
QY 266 CTGCAAGCGGCGCTGCGCTTGGCCCGGGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 325
DB 252 CTGCAAGCGGCGCTGCGCTTGGCCCGGGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 311
QY 326 TTTGCTTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 385
DB 312 TGTGCTTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 371
QY 386 GCGCTGCTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 445
DB 372 GCGCTGCTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 431
QY 446 GTCAAGCGGCGCTGCGCTTGGCCCGGGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 505
DB 432 GTCAAGCGGCGCTGCGCTTGGCCCGGGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 491
QY 506 GCGCTGCTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 565
DB 492 GCGCTGCTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 551
QY 566 GCGCTGCTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 625
DB 552 GCGCTGCTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 611
QY 626 CCAAGTGGCGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 685
DB 612 CCAAGTGGCGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 671
QY 686 GCAAGTGGCGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 745
DB 672 GCAAGTGGCGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 731
QY 746 CCGCTGCTTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 803
DB 732 CCGCTGCTTAAGTGGAGTGGTGAAGCCGCTGCGCTTGGCCCGGGAAGCCGCTGTAACCA 789

```

RESULT 4  
 AAV15072  
 ID AAV15072 standard; DNA; 922 BP.

```

XX AAV15072:
AC 17-OCT-2003 (revised)
XX 19-MAY-1998 (first entry)
DT 19-MAY-1998 (first entry)
XX Hybrid DNA comprising a family 45 cellulase core region.
XX Cellulase core region; isolation; microorganism; identification;
XX hybrid DNA; ss.
XX Humicola nigrescens.
OS Humicola insolens.
OS Chimeric.
XX Key Location/Qualifiers
XX CDS 1..921
XX /tag= a
XX /product= "hybrid gene construct"
XX MO974309-A2.
XX 20-NOV-1997.
XX 12-MAY-1997; 97MO-DX00216.
XX 10-MAY-1996; 96DK-00000562.
XX (NOVO ) NOVO-NORDISK AS.
XX Dalboge H, Diderichsen B, Sandal T, Kauppinen S;
XX MPI; 1998-008878/01.
XX F-PSDB; AAM44259.
XX Isolating novel DNA sequences from microorganisms - without the need for
XX culturing the microorganism.
XX Example 3; Page 51-52; 72pp; English.
XX PS The present sequence represents a novel hybrid gene construct from an
XX CC example of the present invention. The present invention describes a novel
XX CC method for providing a novel DNA sequence encoding a polypeptide from a
XX CC microorganism with an activity of interest. The method comprises: (i) PCR
XX CC amplification of the DNA with PCR primers with homology to (a) known
XX CC gene(s) encoding a polypeptide with an activity of interest; (ii) linking
XX CC the obtained PCR product of a 5' structural gene sequence and a 3'
XX CC structural gene sequence; (iii) expressing the resulting hybrid DNA
XX CC sequence; (iv) screening for hybrid DNA sequences encoding a polypeptide
XX CC with the activity of interest or a related activity; and (v) isolating
XX CC the hybrid DNA sequence identified in step (iv). This method provides for
XX CC identification and isolation of sequences from microorganisms without
XX CC having to cultivate and isolate the microorganism. (Updated on 17-OCT-
XX CC 2003 to standardise OS field)
XX CC
XX Sequence 922 BP; 151 A; 333 C; 262 G; 176 T; 0 U; 0 Other;
XX
XX Query Match 39.5%; Score 451.6; DB 2; Length 922;
XX Best Local Similarity 71.6%; Pred. No. 8.5e-73;
XX Matches 630; Conservative 0; Mismatches 229; Indels 21; Gaps 2;
QY 102 CTTCCTTGGTGGCTTGGCGGCGAGTGGCAAGTCCACAGATATCTGGGACTGC 161
DB 43 CTTCAGGTGGCGGACACTGCTTTTCCTGCTGATGGCAAGTCCAGCGGACTGGGATTC 102
QY 162 TCGAAGCGCTGCTGGGCTGGGCGGAGGCGCGCTCAGCCAAACCGGCTTAAGCGCTGC 221
DB 103 TGTAAAGCGTGGTGGCTGGGCGGAGGCGCTGCTGGAACAGCGCGCTTAAGCGCGCG 162
QY 222 GATGCAACTTCAGCGCTGCTGCACTTCAATGTCCAGTGGGCTGCAAGCGCGCTGC 281
DB 163 AAGCAAAATTCAGCGGATCACCGAACCCCAAGCGCAAGTGGCGGCTGCAGCGGCTTC 222

```





```

Db      659 GCACCGATGCGCGGCAACGACGCGACCTCCCTGCGCTGAGATCCCTCCGCA 718
Qy      772 GCACGCG-----TGGCAGCGGAGCGCCAGCTGAGCTGCGCTGAGGT 813
Db      719 GCGACAGCTCTCCGCTCAACCAAGCTTCCGACGACGACGACGACCTTCCACACT 778
Qy      814 CGGGCCAGAC--GTCTCCCGGCGGCGGAGTGGCTGCACTGCTCAAGAAATGGGCTAGT 870
Db      779 CGAGCCCGCAGTCCAGCTCAAGACTCCGAGCGGCTGCACTGCTGAGAGTGGGCTAGT 838
Qy      871 GCGGTGGATCGGCTTCAAGCGGATGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
Db      839 GCGGCGGCAATGCTGAGCGGCTGCAACGACCTGCTGCTGCTGCTGCTGCTGCTGCT 898
Qy      931 TGAACGACTACTGCTGAGTGGCTCTTAAC 961
Db      899 TTAATGACTGTACCATGAGTGGCTGTAAGC 929

```

## RESULT 6

AAQ26405  
ID AAQ26435 standard; cDNA; 1058 BP.

AAQ26405;

24-OCT-2003 (revised)

25-MAR-2003 (revised)

11-JAN-1993 (first entry)

Humicola insolens DSM 1800 endoglucanase.

Detergent; washing powder; cellulase; softening clay; Cl4CMC-method; flocculation; radio-labelled carboxymethylcellulase; ss.

Humicola insolens; DSM 1800.

Key Location/Qualifiers

FT seq\_peptide

FT mat\_peptide

EP495258-AL.

22-JUL-1992.

06-NOV-1991; 91BP-30202880.

16-JAN-1991; 91BP-30870006.

(PROC) PROCTER &amp; GAMBLE CO.

Convents AC, Busch A, Baek AC;

WPI; 1992-243163/30.

P-PSDB; AAR25525.

Compact, granular detergent compns. - contain high active cellulase

PT and softening clay to provide synergistic effect in softening

performance.

Claim 10; Page 27-28; 38pp; English.

The endoglucanase isolated from H.insolens DSM 1800 has Mr 43,000 and

high cellulase activity (i.e. removes a minimum of 1% of radiolabelled

carboxymethylcellulose (Cl4CMC). The enzyme can be produced

recombinantly using the cDNA coding sequence and used in detergent

compositions along with a surface active agent, a builder system and a

softening clay. See also AAQ26407. (Updated on 25-MAR-2003 to correct PN

field.) (Updated on 24-OCT-2003 to standardise OS field)

Sequence 1058 BP; 189 A; 376 C; 288 G; 205 T; 0 U; 0 Other;

Query Match 37.3%; Score 437.4; DB 2; Length 1058;  
Best Local Similarity 70.4%; Pred. No. 3.3e-70;  
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

```

Qy      55 CCGAGATGCGCTTACCTCCGCTTCTTGACAGACCTTGGCCGCTGCACTTCTGCTG 114
Db      5 CCGAGATGCGCTTACCTCCGCTTCTTGACAGACCTTGGCCGCTGCACTTCTGCTG 64
Qy      115 CCGCGCGCGAGTGGAGGAGTGGCGAGTCCAGAGATAGTGGAGTGGAGTGGAGTGGAGT 174
Db      65 CCGTGGCGCGAGTGGAGGAGTGGCGAGTCCAGAGATAGTGGAGTGGAGTGGAGTGGAGT 118
Qy      175 GCGCTTGGCGCGGAGGAGTGGCGAGTCCAGAGATAGTGGAGTGGAGTGGAGTGGAGT 234
Db      119 GCGGCTGGCGCGAGGAGTGGCGAGTCCAGAGATAGTGGAGTGGAGTGGAGTGGAGT 178
Qy      235 AGGCGCTGCGAGTCCAGATGATGTCAGTGGCGTGGAGTGGAGTGGAGTGGAGTGGAGT 291
Db      179 AGGCTATCGAGACTTCAAGCGCGAGTCCAGAGATAGTGGAGTGGAGTGGAGTGGAGT 238
Qy      292 GCGCGGACGAGTCCCGTGGCGGAGTGGAGTCCAGAGATAGTGGAGTGGAGTGGAGTGGAGT 351
Db      239 GCGCGGACGAGTCCCGTGGCGGAGTGGAGTCCAGAGATAGTGGAGTGGAGTGGAGTGGAGT 298
Qy      352 GCGTGGCGCGGAGTCCCGTGGCGGAGTGGAGTCCAGAGATAGTGGAGTGGAGTGGAGTGGAGT 411
Db      299 GCGTGGCGCGGAGTCCCGTGGCGGAGTGGAGTCCAGAGATAGTGGAGTGGAGTGGAGTGGAGT 358
Qy      412 CCGGTCGCGCGAGGAGTGGAGTCCAGAGATAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 471
Db      359 CCGGTCGCGCGAGGAGTGGAGTCCAGAGATAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 418
Qy      472 GAGTATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 531
Db      419 GAGTATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 478
Qy      532 GCTCGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGT 591
Db      479 GCTCGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGT 538
Qy      592 GCGATTCCTTCCCGCGCGCGCTCAAGCGCGCTGCGCGAGTGGCGGAGTGGCGGAGTGGCGGAGT 651
Db      539 GCGATTCCTTCCCGCGCGCGCTCAAGCGCGCTGCGCGAGTGGCGGAGTGGCGGAGTGGCGGAGT 598
Qy      652 ACGCCGAGCAACCGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 711
Db      599 ACGCCGAGCAATCCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 658
Qy      712 GCTCGGCTGGAGGAGTGGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGT 771
Db      659 GCGACGAGATGCGCGGAGTGGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGTGGCGGAGT 718
Qy      772 GCACGCG-----TGGCAGCGGAGCGCCAGCTGAGCTGCGCTGAGGT 813
Db      719 GCGACAGCTCTCCGCTCAACCAAGCTTCCGACGACGACGACGACGACCTTCCACACT 778
Qy      814 CGGGCCAGAC--GTCTCCCGGCGGCGGAGTGGGCTGCACTGCTCAAGAAATGGGCTAGT 870
Db      779 CGAGCCCGCAGTCCAGCTCAAGACTCCGAGCGGCTGCACTGCTGAGAGTGGGCTAGT 838
Qy      871 GCGGTGGATCGGCTTCAAGCGGATGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
Db      839 GCGGCGGCAATGCTGAGCGGCTGCAACGACCTGCTGCTGCTGCTGCTGCTGCTGCT 898
Qy      931 TGAACGACTACTGCTGAGTGGCTCTTAAC 961
Db      899 TTAATGACTGTACCATGAGTGGCTGTAAGC 929

```

## RESULT 7

AAQ14856  
ID AAQ14856 standard; DNA; 1060 BP.

```

XX AC AA014856;
XX XX
XX DT 25-MAR-2003 (revised)
XX DT 18-FEB-1992 (first entry)
XX XX
XX DE Humicola insolens DSM 1800 endoglucanase.
XX XX
XX KW cellulase; ss.
XX OS Humicola insolens.
XX XX
XX FH Key Location/Qualifiers
XX FT sig_peptide 10..72
XX FT /*tag= a
XX FT 73..924
XX FT mat_peptide /*tag= b
XX XX
XX EN W09117243-A.
XX PD 14-NOV-1991.
XX PF 09-MAY-1990; 90DK-00001159.
XX PR 09-MAY-1990; 90DK-00001159.
XX PR 22-APR-1991; 91DK-00000736.
XX PA (NOVO ) NOVO-NORDISK AS.
XX XX
XX PI Rasmussen G., Mikkelsen JW, Schulein W, Packar SA, Hagen F;
XX PI Hjort C, Hastrup S;
XX DR WPI; 1991-353765/48.
XX DR P-PSDB; AAR15271.
XX XX
XX PT Cellulase prep. comprising endoglucanase enzyme - used in detergents for
XX PT cellulose-contg. fabrics or to improve drainage of paper pulp.
XX PS Claim 11; Page 48; 67pp; English.
XX XX
XX CC The cellulase (i.e. endoglucanase) gene was isolated from a H. insolens
XX CC cDNA library by screening with a probe based on the amino acid sequence
XX CC of the purified enzyme. (updated on 25-MAR-2003 to correct PA field.)
XX CC
XX SQ Sequence 1060 BP; 190 A; 377 C; 288 G; 205 T; 0 U; 0 Other;
XX
Query Match 37.3%; Score 437.4; DB 2; Length 1060;
Best Local Similarity 70.4%; Pred. No. 3.3e-70;
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;
QY 55 CCAAGATGCGCTTACTCCGTTCTTCGCAACCTGCGCGCTGCACTTCTCTGTGCG 114
DB 5 CCAAGATGCGTTCCCTCCCTCCCTCCGTCGCGCTGCGCGCTGCGCGCTGCG 64
QY 115 CTTCCCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 174
DB 65 CTTCCCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 118
QY 175 GCGCTTGGCCGCGGAGGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 234
DB 119 GCGCTTGGCGCGGAGGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 178
QY 235 AGCGCTGTCCGCTTCAATGTCAGTCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 291
DB 179 AGCGTATTCAGGCTTCAATGTCAGTCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 238
QY 292 GCGCGCGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 351
DB 239 GCGCGCGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 298
QY 352 GCATTCGCGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 411
DB 299 CTATTCGCGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 358

```

```

QY 412 CCGGTCCCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 471
DB 359 CCGGTCCCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 418
QY 472 GAAGTAACAGTTCCATATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 531
DB 419 GAGCAACAGCTTCCATATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 478
QY 532 GCTCCAGTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 591
DB 479 CTCCCAAGTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 538
QY 592 GCGATTCCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 651
DB 533 GCGATTCCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 598
QY 652 AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 711
DB 599 AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 658
QY 712 GCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 771
DB 659 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 718
QY 772 GCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 831
DB 719 GCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 778
QY 814 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 870
DB 779 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 838
QY 871 GCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 930
DB 839 GCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 898
QY 931 TGAAGACTACTCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 961
DB 899 TGAAGACTACTCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 929

RESULT 8
AAQ25932
ID AAQ25932 standard; DNA; 1060 BP.
XX AC AAQ25932;
XX AC
XX DT 27-AUG-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 15-JAN-1993 (first entry)
XX XX
XX DE Cellulase contained in a detergent compsn.
XX KW Endoglucanase; immunoreactive; Humicola insolens; cleaning; softening; ss.
XX OS Humicola insolens.
XX XX
XX FH Key Location/Qualifiers
XX FT CDS 9..927
XX FT /*tag= a
XX XX
XX EN E0495554-A1.
XX PD 22-JUL-1992.
XX PF 15-JAN-1992; 92EP-00200101.
XX PR 16-JAN-1991; 91EP-00870006.
XX PR 06-NOV-1991; 91EP-00202881.
XX PA (PROC ) PROCTER & GAMBLE CO.
XX XX

```







Db 659 GCACCGATGCGCGGACGACGAGCACTTCCTCCGCTCCAGATCCCTCCAGCA 718  
Qy 772 GCACCG-----TGGACCGGAGCGCCACGTCGATGCTGCGCTGGGT 813  
Db 719 GCACCGATGCTCCGCTCAACGACGCTACGACGACGACGACGCTCCACCTCCAGCACT 778  
Qy 814 CGGCGCAGAC---GTCTCCGCGCGCGGAGTGGCTGACGCTTCGACAGATGGGCTCAGT 870  
Db 779 CGAGCCGCGCAGTCCGACGCTTACGACTCCGAGCGGCTGACGCTGAGAGGATGGCTCAGT 838  
Qy 871 GCGGTGATGCGCTTACGCGGATGACGACGACGCTGTGTCTGTGACGACGCTGCGAGAGT 930  
Db 839 GCGGCGGAGATGCTGGAGTGGCTGACGACGCTGCGTGGCTGGAGACACTTGGACGAGAGA 898  
Qy 931 TGAACGACTACTACTCGAGTGCCTTAAAC 961  
Db 899 TTAATGACTGTACCAATCAGTGCCTGTAGAC 929

## RESULT 12

AAQ41732

ID AAQ41732 standard; DNA; 1060 BP.

AC AAQ41732;

Df 25-MAR-2003 (revised)

Df 25-AUG-1993 (first entry)

Df Dye transfer inhibiting compsn. cellulase gene.

Df Detergent; homogeneous endoglucanase component; 43kd cellulase; ss.

Df Humicola insolens.

Df Key Location/Qualifiers

Df CDS 10..927

Df sig\_peptide /\*tag= a

Df mat\_peptide /\*tag= b

Df /\*tag= c

Df EP540784-Al.

Df 12-MAY-1993.

Df 06-NOV-1991; 91EP-00202882.

Df 06-NOV-1991; 91EP-00202882.

Df (PROC) PROCTER &amp; GAMBLE CO.

Df Mccorquodale F, Busch A;

Df WPI; 1993-153868/19.

Df P-PSDB; AAR37150.

Df Compn. for inhibiting dye transfer during fabric washing - contains

Df peroxidase, hydrogen peroxide, substrate and cellulase, esp. endo-

Df glucanase from Humicola insolens.

Df Disclosure; Page 16-19; 28pp; English.

Df The sequence is that encoding a cellulase which is characterised in that

Df the cellulase provides at least 10% removal of immobilised radioactive

Df labelled carboxymethyl cellulose according to the ClAQC method at 25 x

Df 10(-6) by weight of the cellulase protein in the test solution. It can

Df be used as part of a compsn. for inhibiting dye transfer. (Updated on 25-

Df MAR-2003 to correct PN field.)

Df Sequence 1060 BP; 190 A; 377 C; 288 G; 205 T; 0 U; 0 Other;

Df Query Match 37.3%; Score 437.4; DB 2; Length 1060;

Best Local Similarity 70.4%; Pred. No. 3.3e-70;

Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

Qy 55 CCAGATGCGCTACTACCTCCGTTCTTTCGACACACCTGAGCGCTGACCTTCTCTGCTG 114

Db 5 CCAAGATGCGTTCTCCCTCCCTCTCTCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 64

Qy 115 CTTCCGCGCAGATGCGAGTGGCCAGTCCAGAGATGCTGGAGCTGTGCAAGCCGCTGCT 174

Db 65 CCTTCGCGCTGATGCGAG-----GTCCACCGCTACTGGAGCTTCCTGCAAGGCTTCGT 118

Qy 125 GCGCTTGGCGCGGAGAGCGCGCTGAGCACAACGCTCTAAGCGTGGAGTGCAGAACTCC 234

Db 119 GCGCTGCGCGCAAGAGGCTCCGTAACACACCTGCTTTTCTTCCAGACGCACTTCC 178

Qy 235 AGCGCTGCTCCGACTTCAATGCTCAAGTGGGCTGCA---CGCGGCTGCTGCTTCTCT 291

Db 179 AGCGTATACAGAGACTTCCAGCGCAAGTCCGAGTGGAGGCGCGGCTGCTGCTGCTG 238

Qy 292 GCGCGCAGCAGACTCCCTGCGCGGTGAAACGACATCTGCTTACGCTTGGCGCGCAG 351

Db 229 GCGCGCAGCAGACCCCATGGGCTGTGAAACGACATCTGCGCTGCTTGGCTGCTG 298

Qy 352 GCATGCGCGCGGCGGATCCTGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 411

Db 299 CTATGCGCGCGCAGATGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358

Qy 412 CCGGCTCCGCTGCGCGGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 471

Db 359 CCGGCTCTGCTGCTGCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 418

Qy 472 GAGGTACCAAGTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 531

Db 419 GCAACCAACCACTTGTATCTCAATCTCCCGGCGGCGGCTGCTGCTGCTGCTGCTG 478

Qy 532 GCTGCGAGTTGGCGGCGCTCCCGCGGCTCAATGAGGCGGCTGCTGCTGCTGCTGCTG 591

Db 479 CTCCCGAGTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 538

Qy 592 GCGATTCCTTCCCGCGCGCTGCAAGCTCCCGGCTGCGGCTGCTGCTGCTGCTGCTG 651

Db 539 GCGATCGGTTCCCGCGCGCTGCAAGCTCCCGGCTGCTGCTGCTGCTGCTGCTGCTG 598

Qy 652 ACGCGCAGACCCGACGCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711

Db 599 ACGCGCAGACCCGACGCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 658

Qy 712 GCTCCGCGCTGCAAGGCGCAAGCACTCCAGCTTCCCGCTTCAACCCCGCAAGCGGTG 771

Db 659 GCACCGATGCGCGCGGACGACGAGCGCACTTCCGCGCTGCAAGTCCCTCCAGCA 718

Qy 772 GCACCG-----TGGACCGGAGCGCCACGTCGATGCTGCTGCTGCTGCTGCTG 813

Db 719 GCACCGATGCTCCGCTCAACGACGCTACGACGACGACGACGACGCTCCACGACCT 778

Qy 814 CGGCGCAGAC---GTCTCCGCGCGCGGAGTGGCTGACGCTTCAAGATGGGCTCAGT 870

Db 779 CGAGCCGCGCAGTCCGACGCTTACGACTCCGAGCGGTGACGCTGAGAGGAGGCTCAGT 838

Qy 871 GCGGTGATGCGGCTTACGCGGATGACGACGCTGTGTCTGAGCAACGCTGCGAGAGT 930

Db 839 GCGGCGGAGATGCTGGAGGCGGCTGACCACTGCTGCTGCGAGCACTTGGACGAGAGA 898

Qy 931 TGAACGACTACTACTCGAGTGCCTTAAAC 961

Db 899 TTAATGACTGTACCAATCAGTGCCTGTAGAC 929

## RESULT 13

AA110182

ID AA110182 standard; DNA; 1060 BP.

AC AA110182;





XX MO9743409-A2.  
 EN 20-NOV-1997.  
 PD 12-MAY-1997; 97WO-DK000216.  
 EF 10-MAY-1996; 96DK-00000562.  
 PR (NOVO ) NOVO-NORDISK AS.  
 XX Dalboge H, Diderichsen B, Sandal T, Kauppinen S;  
 PI MPI; 1998-008878/01.  
 DR P-PSDB; AAM44266.  
 XX Isolating novel DNA sequences from microorganisms - without the need for  
 PT culturing the microorganism.  
 XX Example 3; Page 47-48; 72pp; English.  
 PS The present sequence encodes a protein with endoglucanase activity used  
 CC in an example of the present invention. The present invention describes a  
 CC novel method for providing a novel DNA sequence encoding a polypeptide  
 CC from a microorganism with an activity of interest. The method comprises:  
 CC (i) PCR amplification of the DNA with PCR primers with homology to (a)  
 CC known gene(s) encoding a polypeptide with an activity of interest; (ii)  
 CC linking the obtained PCR product of a 5' structural gene sequence and a  
 CC 3' structural gene sequence; (iii) expressing the resulting hybrid DNA  
 CC sequence; (iv) screening for hybrid DNA sequences encoding a polypeptide  
 CC with the activity of interest or a related activity; and (v) isolating  
 CC the hybrid DNA sequence identified in step (iv). This method provides for  
 CC identification and isolation of sequences from microorganisms without  
 CC having to cultivate and isolate the microorganism  
 XX  
 SQ Sequence 1060 BP; 190 A; 377 C; 288 G; 205 T; 0 U; 0 Other;  
 Query Match 37.3%; Score 437.4; DB 2; Length 1060;  
 Best local Similarity 70.4%; Pred. No. 3.3e-70;  
 Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 532 GCTGCGAGTTGAGCAGGCTCCCGGCTCAATACAGCGCATTTGTCGCGACAGT 591  
 DB 479 CTCCCGAGTTGAGCGGTCTGCGCGACAGCGTACGCGGAGATCTGCTCCGCAAGAGT 538  
 QY 592 GCGATTCCTTCCCGCGCGCTGTAAGCTCGGCTGCACTGCGGTTTATCTGTTCCAGA 651  
 DB 539 GCGATCGGTTCCCGCAAGCGCTCAAGCGCGGCTGCTAGTCCCTTTCAGCTGTTCCAGA 598  
 QY 652 ACGCCCAACACCGCAAGCTTCAAGTCAAGAGTACAGTCCCGCGAGATGTTGCC 711  
 DB 599 ACGCGCAATCCGAGCTTCAAGTCTCCGAGTCAAGTCCAGAGTCCCGAGTCTGCTC 658  
 QY 712 GCTCGGCTGCAAGGCAAGCAAGCTTCAAGTCCCGGCTTCAAGTCCCGAGTCTGCTC 771  
 DB 659 GCAAGCGAGTCCCGCGCAAGCAAGCAAGTCCCGGCTTCAAGTCCCGAGTCTGCTC 718  
 QY 772 GCAAGCG-----TGGCAAGCGAGCGCTTCAAGTCCCGAGTCTGCTC 813  
 DB 719 GCAAGCGAGTCCCGGCTTCAAGCGAGTCCCGAGTCCCGAGTCTGCTC 778  
 QY 814 CGGCTCAGAC---GTCTCCCGCGCGCGCTTCAAGTCCCGAGTCTGCTC 870  
 DB 779 CGAGCGCGAGTCCAGCTTCAAGTCCCGAGTCCCGAGTCTGCTC 838  
 QY 871 GCGGCGAGTCCAGCTTCAAGTCCCGAGTCCCGAGTCCCGAGTCTGCTC 930  
 DB 839 GCGGCGAGTCCAGCTTCAAGTCCCGAGTCCCGAGTCCCGAGTCTGCTC 898  
 QY 931 TGAAGCACTACTCTGAGTCCCGAGTCCCGAGTCCCGAGTCTGCTC 961  
 DB 899 TTAATGACTGTACATGAGTCCCGAGTCCCGAGTCCCGAGTCTGCTC 929

RESULT 15  
 AAV16102  
 ID AAV16102 standard; cDNA; 1060 BP.  
 AC AAV16102;  
 XX 21-JUL-1998 (first entry)  
 DT  
 DE Humicola insolens surface-active endoglucanase gene.  
 XX  
 KW endoglucanase; surface-active; cellulase; detergent; cleaning;  
 XX performance; stain removal; softening; feel; colour; ss.  
 OS Humicola insolens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 10..927  
 FT /tag= a  
 FT /product= "endoglucanase"  
 FT sig\_peptide 10..72  
 FT /tag= c  
 FT mat\_peptide 73..924  
 FT /tag= b  
 XX  
 PK WO9804663-A1.  
 XX  
 PD 05-FEB-1998.  
 PF 25-JUL-1997; 97WO-US013194.  
 PR 30-JUL-1996; 96US-0023125P.  
 PA (PROC ) PROCTER & GAMBLE CO.  
 PI Boyer SL;  
 XX  
 DR MPI; 1998-130664/12.  
 DR P-PSDB; AAM46616.  
 XX Detergent composition containing both surface-active and non-surface







QY 541 TCAGGCGGCTCCCGCGCTCAATACGAGCGGATTTGTCGCGCAACAGTGGATTCTT 600  
DB 541 TCAGGCGGCTCCCGCGCTCAATACGAGCGGATTTGTCGCGCGCAACAGTGGATTCTT 600  
QY 601 TCCCCGCGCTCAAGCGCGCTGCGCATGCGGTTGATCTGTTCCGAAAGCGCGACA 660  
DB 601 TCCCCGCGCTCAAGCGCGCTGCGCATGCGGTTGATCTGTTCCGAAAGCGCGACA 660  
QY 661 ACCGAGCTTCAAGTTCAGAGGTCAGTGCCTCCCGCGAGATGTTGCGCGCTCCGCT 720  
DB 661 ACCGAGCTTCAAGTTCAGAGGTCAGTGCCTCCCGCGAGATGTTGCGCGCTCCGCT 720  
QY 721 CGAAGCGCAACAGCACTTCAGCTTCCCGCTTTTCAACCCCGCAAGCGGTGCAACGCTG 780  
DB 721 CGAAGCGCAACAGCACTTCAGCTTCCCGCTTTTCAACCCCGCAAGCGGTGCAACGCTG 780  
QY 781 GCACCGGAGCGCCAGCTGCACTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCA 840  
DB 781 GCACCGGAGCGCCAGCTGCACTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCA 840  
QY 841 GTGCGTCACTGCTCAAGAGTGGGCTCAGTGGCGGTGCGATCGGCTTCAAGCGATGCA 900  
DB 841 GTGCGTCACTGCTCAAGAGTGGGCTCAGTGGCGGTGCGATCGGCTTCAAGCGATGCA 900  
QY 901 CCGTGTCTCTGCGCACTGCGCGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 960  
DB 901 CCGTGTCTCTGCGCACTGCGCGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 960  
QY 961 CAGCTTTTCCGCAAGTGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 1020  
DB 961 CAGCTTTTCCGCAAGTGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 1020  
QY 1021 TTTTGAAGCGCTCAATACATATACCTTCATCTGTCATAGCAAGCGGTACCA 1080  
DB 1021 TTTTGAAGCGCTCAATACATATACCTTCATCTGTCATAGCAAGCGGTACCA 1080  
QY 1081 TCTCAACCGCACTTTGAGGAGCGGAGTCAAGCGGCTTTTAAAAA 1140  
DB 1081 TCTCAACCGCACTTTGAGGAGCGGAGTCAAGCGGCTTTTAAAAA 1140  
QY 1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174  
DB 1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174

## RESULT 3

US-09-229-911A-11

Sequence 11, Application US/09229911A

Patent No. 6387690

GENERAL INFORMATION:

APPLICANT: Schuelein, Martin

Andersen, Lene N.

Lassen, Soren F.

Kauppinen, Markus S.

Lange, Lene

Nielsen, Ruby I.

Ihara, Michiko

Takagi, Shinobu

TITLE OF INVENTION: No. 6387690e1 Endoglucanases

NUMBER OF SEQUENCES: 109

CORRESPONDENCE ADDRESS:

ADDRESS: No. 6387690 No. 6387690disk of No. 6387690th America, Inc.

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/229,911A

FILING DATE: 13-Jan-1999

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/651,136

FILING DATE: 21-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4366-200-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 1174 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 60..956

SEQUENCE DESCRIPTION: SEQ ID NO: 11:

US-09-229-911A-11

Query Match

Best Local Similarity 100.0%; Score 1174; DB 4; Length 1174;

Best Local Similarity 100.0%; Pred. No. 1.9e-225;

Matches 1174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAGACCCCTTCAAGCTTCAAGTTTCACCCCGCTCTTTTCTTGGCCCCCGAGA 60  
DB 1 GAGGAGACCCCTTCAAGCTTCAAGTTTCACCCCGCTCTTTTCTTGGCCCCCGAGA 60  
QY 61 TGGGCTCTACCTCGGTTCTTTCGCACAACCTTGGCCCTGCACTTCTCTGTTCCG 120  
DB 61 TGGGCTCTACCTCGGTTCTTTCGCACAACCTTGGCCCTGCACTTCTCTGTTCCG 120  
QY 121 CGGCAAGTGGCGGAGTTCAGAGGATCTGCGAAGTCTGCGAAGTCTGCGAAGTCTG 180  
DB 121 CGGCAAGTGGCGGAGTTCAGAGGATCTGCGAAGTCTGCGAAGTCTGCGAAGTCTG 180  
QY 181 GGGCGGAGAGGCGCGCTGAGCAACCGGCTTACGCGTGCAGTGCAGTTCGAGGCG 240  
DB 181 GGGCGGAGAGGCGCGCTGAGCAACCGGCTTACGCGTGCAGTGCAGTTCGAGGCG 240  
QY 241 TGTCCAGCTTCAATGTCAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 300  
DB 241 TGTCCAGCTTCAATGTCAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 300  
QY 301 AGACTCCCTGGGCGGTGAGCAATCTGCTTACGCTTACGCTTACGCTTACGCTTAC 360  
DB 301 AGACTCCCTGGGCGGTGAGCAATCTGCTTACGCTTACGCTTACGCTTACGCTTAC 360  
QY 361 GCGGCTCCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 420  
DB 361 GCGGCTCCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTC 420  
QY 421 TCGCGGCAAGCAATGTCAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 480  
DB 421 TCGCGGCAAGCAATGTCAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 480  
QY 481 AGTTCAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 540  
DB 481 AGTTCAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG 540  
QY 541 TCGCGGCTCCCGCGCTCAATACGCGGATTTGTCGCGCAACAGTGGATTCTT 600  
DB 541 TCGCGGCTCCCGCGCTCAATACGCGGATTTGTCGCGCAACAGTGGATTCTT 600  
QY 601 TCCCCGCGCTCAAGCGCGCTGCGCATGCGGTTGATCTGTTCCGAAAGCGCGACA 660  
DB 601 TCCCCGCGCTCAAGCGCGCTGCGCATGCGGTTGATCTGTTCCGAAAGCGCGACA 660

QY	162	TGGAAGCGGTGTGCGCTTGGCCCGGGAAGCCCGGTGACGAAACCGGTCTAAGGTC	222
Db	103	TGTAAGCGGTGTGCTGTGTGCGCCGGAAGGCGCTGTGAACGACCCGTTACGCCCCG	162
QY	222	GATGCGAAGCTTCCAGAGCGCTGTCCGACTTCATATGTCTAGTGGGCTTGCAAGCGAGCTCG	281
Db	163	AAGCGAAAGCTTCCAGCGCATACCGAACCCCAAGCGCAAGTCCGGCTCGAGATGGGCGCTCC	222
QY	282	GGCTACTCTTCGCGCGACGACACTCCCTTGAGGGTGAACGACATCTCGGCTTAAGGCTTC	341
Db	223	GGCTTCTCTTCGCGCGACGACGCCCTCGGCGCGTGAACGACATTTGCTCTAAGGTTTC	282
QY	342	GCGCGGAGACAGCATCCGCGCGGGGTCCGAATCTGTGTGTGTGTCGCGCTGTACGCGCTC	401
Db	283	GCGGCTAAGCGCGCTCGCGCGCGATTCGAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	342
QY	402	ACCTTCACTTCCGAGTCCCGTTCGCGGACGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGT	461
Db	343	ACCTTCACTTCCGAGTCCCGTTCGAGAAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	402
QY	462	GCGGACCTTGGGAATGAACGATTTGATGTGATGCGATGCGCGCGCGCGCGCTGTGAGGACTTC	521
Db	403	GCGGACCTTGGGAATGAACGATTTGATGTGATGCGATGCGCGCGCGCGCGCTGTGAGGACTTC	462
QY	522	AACGGCTGACGCTGCGAGTTTGCGCGGCTCCCGCGCTCATATGCGCGGACTTTCTCTCG	581
Db	463	GAGCGCTGTGTGCTGACGAGTTTGCGGCTCTCGCGCGGCGCGCGCTGTGTGTGTGTGTGTGT	522
QY	582	CGCGACAGTGTGATCTCTTCCCGCGCGCGCTGTGACGCGCGCGCGCGCGCTGTGAGGACTTC	641
Db	523	CGCGAGAAATGTGACTCTTCTTCCCGCGCGCGCTGTGACGCGCGCGCGCGCTGTGAGGACTTC	582
QY	642	TGGTTCGAAAGCGCGCAACCGCGACGTTTCAGCTTCAGAGGTGTGTGTGTGTGTGTGTGTGTGT	701
Db	583	TGGTTCGAAAGCGCGCAACCGCGACGTTTCAGGCTTCGCGAGGTGTGTGTGTGTGTGTGTGTGT	642
QY	702	ATCGTGTCCCGCTTCGCGTGAAGCGCGACGACGATCCAGTTCGCCCGTTCACGCCCC	761
Db	643	CTCGTGTCCCGCTTCGCGATGCGCGCGACGACGAGTTCGCGCGCTGTGTGTGTGTGTGTGTGT	702
QY	762	CCAGCGGTGTGCGAAAGG-----TGCGACCGGAGCGCGCCAGCTGCGACT	803
Db	703	CCCTTCGACGACGACGACGCTCTCCGGTTCACGACGCTTACGACGACGACGACGACGACGACG	762
QY	804	GCGCGCTGTGCGCGCGCGACG-----GTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	860
Db	763	TCGACGACCTTCGAGCG	822
QY	861	TGGGCTGATGCGGTGTGCGATGCGCTTCAGCGGATGACGACGCTGTGTCTGTGCGACGCG	920
Db	823	TGGGCTGATGCG	882
QY	921	TGCGAAGATGTGAAGGACTTCTACTGCGAGTCTCTTAA 960	
Db	883	TGCGAAGATGTGAAGGACTTCTACTGCGAGTCTCTTAA 922	
RESULT 5			
US-08-090-013-1			
; Sequence 1, Application US/08090013			
; Patent No. 5443750			
GENERAL INFORMATION:			
; APPLICANT: CONVENTS, ANDRE C			
; APPLICANT: BUSCH, ALFRED			
; APPLICANT: BARCK, ANDRE C			
; TITLE OF INVENTION: DETERGENT COMPOSITION WITH HIGH ACTIVITY			
; TITLE OF INVENTION: CELLULASE AND SOFTENING CLAYS			
; NUMBER OF SEQUENCES: 4			
; CORRESPONDENCE ADDRESS:			
; ADDRESS: THE PROCTOR & GAMBLE COMPANY			
; STREET: 5299 SPRING GROVE AVENUE			
; CITY: CINCINNATI			
; STATE: OHIO			

1 COUNTRY: USA  
2 ZIP: 45217  
3  
4 COMPUTER READABLE FORM:  
5 MEDION TYPE: Floppy disk  
6 COMPUTER: IBM PC compatible  
7 OPERATING SYSTEM: PC-DOS/MS-DOS  
8 SOFTWARE: Patent In Release #1.0, Version #1.25  
9  
10 CURRENT APPLICATION DATA:  
11 APPLICATION NUMBER: US/08/050,013  
12  
13 FILING DATE:  
14 CLASSIFICATION: 435  
15 PRIOR APPLICATION DATA:  
16 APPLICATION NUMBER: EP 9120280.0  
17 FILING DATE: 06-NOV-1991  
18 ATTORNEY/AGENT INFORMATION:  
19 NAME: PATEL, KEN K.  
20 REGISTRATION NUMBER: 33,988  
21 REFERENCE/DOCKET NUMBER: CM393  
22 TELECOMMUNICATION INFORMATION:  
23 TELEPHONE: 513-627-6437  
24 TELEFAX: 513-627-4854  
25  
26 INFORMATION FOR SEQ. ID NO.: 1:  
27 SEQUENCE CHARACTERISTICS:  
28 LENGTH: 1060 base pairs  
29 TYPE: nucleic acid  
30 STRANDEDNESS: unknown  
31 TOPOLOGY: unknown  
32  
33 FEATURE:  
34 NAME/KEY: CDS  
35 LOCATION: 10..924  
36  
37 US-08-090-013-1

Query	Match	Simlarity	Score	DB 1	Length
37.33	Score 437.4	DB 1	Length 1060		
Best Local Similarity 70.4%	Pred. No. 1.3e-78				
Matches 653	Conservative 0	Mismatches 246	Indels 30	Gaps 4	
QY	55	CCAGATGAGGCTTACATCCCGTTCCTTGCCAAACCCCTGACGACTTCTCTGATCG	114		
Db	5	CCAAATATGATTTCTCTCCCTCTCTCCGCTGCGCGTGTGAGCGGCTCGCGGATTTGG	64		
QY	115	CTTCGCGGCGCAATGCGAGTGGCCAGTGCACAGATATCTGGAGACTGCTGCAGAGCCGTGCT	174		
Db	63	CCCTTCACGCTGATGAGGAG-----GTCCACCGGCTACTCGGAGCTGCTGCAGAGCTTCGT	118		
QY	175	GCCCTTGAGCCGCGGAGAGGCGCGGTCAGGCAACGCGTCTACGCGTGTGATGCAACTTCC	234		
Db	119	GCGGCTGGGCGCAAGAGAGCTCCCGTGAACAGGCTGTCTTTTCTTGAAAGCGCACTTCC	178		
QY	235	AGGAGCTGTGCGGACTTTCATATTCACAGTGGGACTGCA---CGAGCGCTCGACTACTCT	291		
Db	179	AGGGTATACAGGACTTTCAGACGACAGTCCGAGCTGGAGGCGGAGTGTGCTACTCTGT	238		
QY	232	GCGCGGACCAAGCTCTCTGGGCGGTGAACGACATTTGGCTTAAGCTTGTGCGTACACA	351		
Db	239	GCGCGGACCAAGCCCCATGGGCTGTGAACACACACTTGGCGCTGGTTTGTATCCACT	298		
QY	352	GCAATGCGCGGAGGGTCCGAAATCTGTGAGTGTGCGGCGCTGATACGCGCTCACTTCACTT	411		
Db	239	CTATTTCCCGGACGCAATAGAGCGGCTGTGTCTGCGCTGTCTACAGCTACACTTTCACAT	358		
QY	412	CCGAGTCCCGTGGCGCGGACAGACATGATGTGTGACGTCAACGACGACTGGGCGGACCTGG	471		
Db	359	CCGAGTCTGTGTGTGTGAGAAAGATGGTGTGTCAAGTCCACAGCACTGGCGGTAATCTTG	418		
QY	472	GAAGTAAACAGTTGATATGCGATATGCCATGCCCGGCGGCGGTGGGATCTTTCAGCGGCTGCA	531		
Db	419	GCGAGCAACCACTTGAATCTCAACATGCCCGGCGGCGGTGTGGATATCTTGAAGGATGCA	478		
QY	532	GCTGCGAGTTGGGCGGCGCTCCCGGCGGCTCAATTAACGCGCGGATTTGGTGTGCGCGACAGT	591		
Db	479	CTCCCCAGTGGGCGGCTGTGCCCGGACAGGCGCTACGGGCGGATGTGTGTCCGCAACAGT	538		
QY	592	GCGATCTCTCCCGCGCGGCTCAAGCCCGGCTGCCAGTGGGCGTTTGACTGTTCAGCA	651		

Db	539	GCATTCGGTTCCTCCGACGCGCTTCAGCCCGCTGCTACTGCGGCTTGCATCGGTTTCAGA	596
Qy	652	ACGCCACAAACCAGAGCTTCAGGTTTCAGACAGGTGCATGTCCTCCGCGAGATCTGTGCC	711
Db	599	ACGCCACAAATCCAGGCTTCAGCTTCCTCAGGCTTCAGATGCCAGCCAGACTGTGTGCTC	658
Qy	712	GCTCCGGCTGCAAGGGCAACGACATCAGACTTCCCGCTTCACCCCCCAAGGGG	777
Db	659	GCACCGAATGCCGCGCGACAGACAGGCAACGGCACTTCTTCGTCTCCAAATCTCCCTCAGCA	718
Qy	772	GCACCG-----TGCACCGGACGCGCCACGTCAGTGGCGCTGGGT	813
Db	719	GCACAGCTCTCCGGTCAACCAAGCTTACAGACACAGACACACGTCACCTCCACCACT	778
Qy	814	CGGCGCCAGC---GTCTCCCGCGCGGCGGCGAGTGTGCGACGTCACAAATGGGCTCACT	877
Db	779	CGAGCCCGCAGCTCCAGCTTCAGACTCCCAACCGGCTGCGATGCTGAAAGGTGGACTCACT	838
Qy	871	CGGCTGGCATCGGCTTCAGCGGATGCACACCACTGTGTCTTGGCACAACCTGCCAGAAAT	930
Db	839	CGCGCGGCAATGGCTGGAAGGGGTGCAACACTGCGTCCGTGGCAGCACTTGACAGAAAGA	898
Qy	931	TGAACGACTACACTCCCAATGCGCTTAAAC	961
Db	899	TTAATGACTGGTACCATCATGATGCTGTGATC	929

RESULT 6  
 US-08-081-328-1  
 Sequence 1, Application US/08081328  
 Patent No. 5520838  
 GENERAL INFORMATION:  
 APPLICANT: BAECK, ANDRE C.  
 APPLICANT: CEULEMANS, RAPHAEL ANGELENE A.  
 APPLICANT: BUSCH, ALFRED (NMN)  
 TITLE OF INVENTION: COMPACT DETERGENT COMPOSITIONS WITH  
 TITLE OF INVENTION: HIGH ACTIVITY CELLULASE  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: THE PROCTOR & GAMBLE COMPANY  
 STREET: 1810 East Miami River Road  
 CITY: CINCINNATI  
 STATE: OHIO  
 COUNTRY: USA  
 ZIP: 45253-8707  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/081,328  
 FILING DATE: 11/19/93  
 CLASSIFICATION: 252  
 ATTORNEY/AGENT INFORMATION:  
 NAME: ZERBY, KIM WILLIAM  
 REGISTRATION NUMBER: 32,323  
 REFERENCE/DOCKET NUMBER: C0356M  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 513-627-2885  
 TELEFAX: 513-627-0318  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1060 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 10..924  
 US-08-081-328-1

Query Match 37.3%; Score 437.4; DB 1; Length 1060;  
Best Local Similarity 70.4%; Pred. No. 1.3e-78;  
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 33 CGAGATGACCTCTACTCCCGCTTCTGCAACCCCTGCGCTGCACTTCTCTGCG 114  
DB 5 CCAAGATGCTTCTCTCCCTCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 64  
QY 115 CTTCCGCGCCGAGTGGCACTGCACTGCAAGATCTGCACTGCTGCAAGCCGCT 174  
DB 65 CTTGCGCGCTGATGCGAG-----GTCCACCGCTACTGCGAGCTGCGAGCCCTG 118  
QY 175 GCGCTTGGCCCGGAGAGCGCGCTGCAAGCCGCTGCAAGCTGCAAGCTGCA 234  
DB 119 GCGCTTGGCCCGAGAGAGCTCCCGTGAACGCTGCTCTTCTGCAAGCGCACTTCC 178  
QY 235 AGCGCTGTCCGACTTCATGTCAGTGGGCTGCA---CGCGGCTGGGCTGCTCT 291  
DB 179 AGCGATACGAGACTTCAGCGCAAGTCCGCTGCGAGCGCGGCTGCTGCTGCT 238  
QY 292 GCGCGACGAGACTCCCTGCGGCTGCAAGCAATGCTGCTGCTGCTGCTGCTG 351  
DB 239 GCGCGACGAGACCCCTGCGGCTGCAAGCAATGCTGCTGCTGCTGCTGCTG 298  
QY 352 GCGCGCGCGCGGCTGCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 411  
DB 299 CTAATGCGCGAGCAATGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 358  
QY 412 CGGCTCCGCTGCGCGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 471  
DB 359 CGGCTCCGCTGCTGCGAGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 418  
QY 472 GAATGATCAGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 531  
DB 419 GCAAGAACCACTTCAGATCAGCAATCCCGCGCGCGCTGCTGCTGCTGCTGCTG 478  
QY 532 GCTGCAAGTTCGCGCGCTCCCGCGCGCTCAATGCTGCTGCTGCTGCTGCTG 591  
DB 479 CTTCCCGATTCGCGCGCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 538  
QY 592 GCGATTCCTTCCTCCCGCGCGCTCAAGCTCCCGCGCTGCTGCTGCTGCTGCTG 651  
DB 539 GCGATCGGCTTCCTCCCGCGCGCTCAAGCTCCCGCGCTGCTGCTGCTGCTGCTG 598  
QY 652 AGCGCGCAAGCGGAGTTCAGATTCAGAGAGTGAATGCTGCTGCTGCTGCTGCTG 711  
DB 599 AGCGCGCAAGTTCGAGCTTCAGATTCAGAGAGTGAATGCTGCTGCTGCTGCTG 658  
QY 712 GCTCGGCTGCAAGCGCAAGCACTGCAAGCTTCCTGCTGCTGCTGCTGCTGCTG 771  
DB 659 GCACTCGATTCGCGCGCAAGCAAGCGCAAGCTTCCTGCTGCTGCTGCTGCTGCTG 718  
QY 772 GCAACG-----TGGCAGCGGAGCGCCCACTGCTGCTGCTGCTGCTGCTG 813  
DB 719 GCAACGATTCCTGCTGCAAGCAAGCTTCAGAGCAAGCAAGCAAGCAAGCTTC 778  
QY 814 CGGCGCAAGC---GTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 870  
DB 779 CGAGCGCGCAAGCTTCAGAGCTTCAGAGCGCGCGCGCGCGCGCGCGCGCGCTG 838  
QY 871 GCGGTCGATTCGCTTCAGCGAGTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGT 930  
DB 839 GCGGTCGATTCGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAGT 896  
QY 931 TGAACGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961  
DB 899 TTAATGACTGATCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 929

## RESULT 7

US-08-232-249-1  
; Sequence 1, Application US/08232249  
; Patent No. 5610129

GENERAL INFORMATION:  
? APPLICANT: MCCORD, FINLAY (MMN)  
? APPLICANT: BUSCH, ALFRED (NNN)  
? TITLE OF INVENTION: DYE TRANSFER INHIBITING COMPOSITIONS  
? NUMBER OF SEQUENCES: 4  
? CORRESPONDENCE ADDRESS:  
? ADDRESS: THE PROCTER & GAMBLE COMPANY  
? STREET: 5299 SPRING GROVE AVENUE  
? CITY: CINCINNATI  
? STATE: OHIO  
? COUNTRY: USA  
? ZIP: 45217  
? COMPUTER READABLE FORM:  
? MEDIUM TYPE: Floppy disk  
? COMPUTER: IBM PC compatible  
? OPERATING SYSTEM: PC-DOS/MS-DOS  
? SOFTWARE: Patentin Release #1.0, Version #1.25  
? CURRENT APPLICATION DATA:  
? APPLICATION NUMBER: US/08/232,249  
? FILING DATE: 05-MAY-1994  
? CLASSIFICATION: 510  
? PRIOR APPLICATION DATA:  
? APPLICATION NUMBER: EP 91202882.6  
? FILING DATE: 06-NOV-91  
? ATTORNEY/AGENT INFORMATION:  
? NAME: ALLEN, GEORGE W.  
? REGISTRATION NUMBER: 26,143  
? REFERENCE/DOCKET NUMBER: CM-395  
? TELECOMMUNICATION INFORMATION:  
? TELEPHONE: 513-627-5946  
? TELEFAX: 513-627-8118  
? INFORMATION FOR SEQ ID NO: 1:  
? SEQUENCE CHARACTERISTICS:  
? LENGTH: 1060 base pairs  
? TYPE: nucleic acid  
? STRAND: BDNES: unknown  
? TOPOLOGY: unknown  
? FEATURE:  
? NAME/KEY: CDS  
? LOCATION: 10..924  
? US-08-232-249-1

Query Match 37.3%; Score 437.4; DB 1; Length 1060;  
Best Local Similarity 70.4%; Pred. No. 1.3e-78;  
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 55 CGAGATGACCTCTACTCCCGCTTCTGCAACCCCTGCGCTGCACTTCTCTGCG 114  
DB 5 CCAAGATGCTTCTCTCCCTCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 64  
QY 115 CTTCCGCGCCGAGTGGCACTGCACTGCAAGATCTGCACTGCTGCAAGCCGCT 174  
DB 65 CTTGCGCGCTGATGCGAG-----GTCCACCGCTACTGCGAGCTGCGAGCCCTG 118  
QY 175 GCGCTTGGCCCGGAGAGCGCGCTGCAAGCCGCTGCAAGCTGCAAGCTGCA 234  
DB 119 GCGCTTGGCCCGAGAGAGCTCCCGTGAACGCTGCTCTTCTGCAAGCGCACTTCC 178  
QY 235 AGCGCTGTCCGACTTCATGTCAGTGGGCTGCA---CGCGGCTGGGCTGCTCT 291  
DB 179 AGCGATACGAGACTTCAGCGCAAGTCCGCTGCGAGCGCGGCTGCTGCTGCTG 238  
QY 292 GCGCGACGAGACTCCCTGCGGCTGCAAGCAATGCTGCTGCTGCTGCTGCTG 351  
DB 239 GCGCGACGAGACCCCTGCGGCTGCAAGCAATGCTGCTGCTGCTGCTGCTG 298  
QY 352 GCGCGCGCGCGGCTGCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 411  
DB 299 CTAATGCGCGAGCAATGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 358  
QY 412 CGGCTCCGCTGCGCGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 471  
DB 359 CGGCTCCGCTGCTGCGAGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 418



[illegible]

RESULT 8  
US-08-921-426-7  
; Sequence 7, Application US/08921426  
; Patent No. 5837847

GENERAL INFORMATION:  
APPLICANT: Royer, John C  
APPLICANT: Moyer, Donna L  
APPLICANT: Yoder, Wendy T  
APPLICANT: Shuster, Jeffrey R  
TITLE OF INVENTION: NON-TOXIC, NON-TOXIGENIC, NON-PATHOGENIC  
TITLE OF INVENTION: FUSARIUM EXPRESSION SYSTEM  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESS: No. 58378470 No. 5837847disk of No. 5837847th America, Inc  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/921,426  
FILING DATE: 29-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/456,433  
FILING DATE: 01-JUN-1995  
APPLICATION NUMBER: US 08/404,678  
FILING DATE: 15-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Agtis Dr., Cheryl H.

```

1 REGISTRATION NUMBER: 34,086
2 REFERENCE/DOCKET NUMBER: 4216.010-US
3 TELECOMMUNICATION INFORMATION:
4 TELEPHONE: 212-867-0123
5 TELEFAX: 212-878-9655
6 INFORMATION FOR SEQ ID NO: 7:
7 SOURCE CHARACTERISTICS:
8 LENGTH: 1060 base pairs
9 TYPE: nucleic acid
10 STRANDEDNESS: single
11 TOPOLOGY: linear
12 FEATURES:
13 NAME/KEY: CDS
14 LOCATION: 10..924
15 FEATURES:
16 NAME/KEY: mat_peptide
17 LOCATION: 73..924
18 FEATURES:
19 NAME/KEY: sig_peptide
20 LOCATION: 10..72
21
22 US-08-921-426-7

```

Query Match	37.3%	Score 437.4	DB 2	Length 1060
Best Local Similarity	70.4%	Seed No. 1,3e-78		
Matches 655	Conservative 0	Mismatches 246	Indels 30	Gaps 4
QY	55	CGAGATGAGGCTTACTCTCCCGTCTTTCGTGGACAACAACCTTGGCCGTGCACTTCTCTGGTGG	114	
Db	5	CCAAAGATGATGTTCTCTCCCGCTCTCCCGCGCGCGTGGTGGCCCGCCCTGGCCGGTGGTGG	64	
QY	115	CCCTCCGCGGACAGTGGACATGGACAGTGCACAGATATCTGGAACCTCTGTCACAAAGCCGTGGT	174	
Db	65	CCCTTTCGCGCTGAATGGGAG-----GTTCACCCCGCTATCTGGAAGCTGCTGACAGCCTTCGT	118	
QY	175	GCGCTTGGCCCGGAGAGGCGCGCTGACGCACACGGTCTGACGCTGACGCTGACGCTGACCTTC	234	
Db	119	GCGGCTGGGGCCAGAAAGAGCGCTCCCTGGAACACGCTGTCTTTTCTTGCAACGCAACTTCC	178	
QY	235	AGGCGCTGTCCGACCTTCATATGTCAGTCCGGGCTTGGAA----GGCGGCTTCGGCTTACCTCT	291	
Db	179	AGGCTATACGAAATTTGATAGCGCAAGTCCGGCTGGACCGCGGCGGTGTGGCTTACTCTGT	238	
QY	292	GCGCGACACAGACTCTCCTGGGCGGTGAACGACATCTTGCTTACGAGCTTTCGCGCGACGA	351	
Db	239	GCGCGACACAGACCCCATGGGCTGTGAACGACATCTTGCGGCTTGGTTTGGCTGACCT	298	
QY	352	GCAATCGCGGCGGCTCCGAAATCTTCGTGGTGGCTGGGCTCTGTACGGCTTACCTTCACTT	411	
Db	299	CTATTGCGCGGAGAAATGAGGCGGCGTGGCTGGCGCTCTCTTCAAGACTCACTTCAACT	358	
QY	412	CCGGTCCCGTTCGCGGACAGACAAATGGTGGTGTCAATGCACAGACACTGGCGGACACTGG	471	
Db	359	CGGGTCTGTTTGGCGCAAGAAATGGTGGTCACTGCACACACACACTGGCGGTGATCTTG	418	
QY	472	GAAATACCAATCTCATATTCGCAATGCGCGCGCGGCGGTGGCACTTTCAACGCGTCGA	531	
Db	419	GCAAGAACATTTCAATCTCAATCAATCCCGGCGGCGGTGGCACTTTCAACGAGATGA	478	
QY	532	GTTCGCAATTCGCGCGGCTCCCGCGGCGTCAATATGCGGCGCAATTTGGTCGCGACAGT	591	
Db	479	CTCCCAATTCGCGGCTGTGCTCCCGCGGCGGTGAGCGGCGCACTGTGTCCGCAACGAT	538	
QY	592	GCGATTCCTTCCCGCGCGCGCTCAAGCTCGGCTGACATGGCGGTTTGAATGTTTCAGA	651	
Db	539	GCGATTCGATTCCTCCGACGCGCTCAAGCGCGGCGGTGCTACTGGCGCTTCACTGGTTTCAAA	598	
QY	652	ACGCGCAACAACCGACGTTCAAGTTCCAGAGGTGACAGTCCCGCGACGATGCGTTTGGCC	711	
Db	599	ACGCGCAACAATCCGACTTCAAGCTTCGCTGAGTGTCAATGCTCCACGACGASCTGGTGGCTC	658	
QY	712	GCTTCGCGCTGACGACGACGACTCGAGCTTCCCGCTTTCACCCCCCGAAGCGGTG	771	
Db	659	GCAACGAGATGCGCGCAACACGACGAGCAATTCCTCGCGTGCACAGATCCCTCCACAGA	718	



GENERAL INFORMATION:  
APPLICANT: SVENDSEN, Allan  
APPLICANT: VON DER OSTEN, Claus  
APPLICANT: CLAUSEN, Ib Groth  
APPLICANT: PATKAR, Shankar Anant  
APPLICANT: BORCH, Kim  
TITLE OF INVENTION: STABILIZED ENZYMES  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESS: No. 59143060 No. 5914306disk of No. 5914306ch America  
STREET: 405 Lexington Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/140,008A  
FILING DATE: 22-OCT-1993  
CLASSIFICATION: 252  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 3601,204-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 867-0123  
TELEFAX: (212) 878-9655  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1060 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 10..924  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 10..72  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 73..924  
US-08-140-008A-3

Query Match 37.3%; Score 437.4; DB 2; Length 1060;  
Best Local Similarity 70.4%; Pred. No. 1.3e-78;  
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 55 CCAGATGCGCTCTACTCCGTTCTTCGCAACCTTGACCGCTGCACTTCTCTGTCG 114  
DB 5 CCAAGATGCGCTCTCTCCCTCTCTCCGCTGCGCGTTGAGCGCGCTGCGCTGTCG 64  
QY 115 CCGCGCGCGCGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 174  
DB 65 CCGTTCGCGCTGATGGAG-----GTCCACCGCTCTCTGAGCTGCTGCAAGCTTCG 118  
QY 175 GCGCTTGCGCCGAGAGGCGCGCTGAGCGCAACCGGCTCTACGCGTGCAGTCCAACTTCG 234  
DB 119 GCGGCTGGGCGCAAGAGGCTCCGCTGACGAGCTCTCTTCTCTGCAAGCGCACTTCG 178  
QY 235 AGGCGCTGTGCGACTTCAATGTCACGCTGCGCTGCA--CGCGCGCTGCGCTTCTCT 291  
DB 179 AGCGTATCACGAGCTTCAAGCGCAAGTCCGAGCTGCGCGCGCGCTGCTGCGCTTCTCT 238  
QY 292 GCGCGAGCGAGCTCCGCGCGCGGAGAGAGCAATCTGCGCTTACGCGCTGCGCGAGCA 351  
DB 239 GCGCGAGCGAGCGCGCGCGCTGAGAGAGAGCACTTCTGCGCTGCGCTTCTGCGAGCT 298

QY 352 GCATGCGCGCGGCTCCGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 411  
DB 299 CATTTCGCGGAGGAGATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358  
QY 412 CCGGTCGCGCGCGGAGAGATGAGTGGGAGTGAAGCACTGCGCGCGAGCGCGAG 471  
DB 359 CCGGTCGCTGCTGCGAGAGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 418  
QY 472 GAAGTACCAAGTTCGATATGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 531  
DB 419 GCAGCAACCACTGATCTCAACATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 478  
QY 532 GCTGCGAGTGGGCGGCTCCCGCGCGCTCAATGCGCGCGCGCGCGCGCGCGCGCGCG 591  
DB 479 CTCCTCGATGGGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 538  
QY 592 GCGATTCCTTCG 651  
DB 539 GCGATGCGTTCGCGAGCGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 598  
QY 652 AGCGCGAAGCGCGAGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 711  
DB 599 AGCGCGAAGCGCGAGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 658  
QY 712 GCTGCGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 771  
DB 659 GCAAGCGAGCGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 718  
QY 772 GCAAGCG-----TGGCACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 813  
DB 719 GCAAGCGCGCTCCCGCTCAACAGCGCTTCAAGCGAGCGAGCGAGCGAGCGAGCGAG 778  
QY 814 CCGCGCGAGC---GTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 870  
DB 779 CGAGCGCGCGAGTCAAGCGAGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 838  
QY 871 GCGGCGAGTGGCTTCAAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 930  
DB 839 GCGCGAGAGTGGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 898  
QY 931 TGAAGACTACTCTGCGAGTGCCTTAAC 961  
DB 899 TTATGACTGTGACATGAGTCCGTGAGC 929

RESULT 11  
US-08-389-423-1  
Sequence 1, Application US/08389423  
Patent No. 5948672  
GENERAL INFORMATION:  
APPLICANT: Rasmussen, Grethe  
APPLICANT: Mikkelson, Jan Moller  
APPLICANT: Schuelein, Martin  
APPLICANT: Patkar, Shankar A.  
APPLICANT: Hagen, Fred  
TITLE OF INVENTION: A Cellulase Preparation Comprising an  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESS: No. 59486720 No. 5948672disk of No. 5948672th America, Inc.  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/389,423  
FILING DATE: 14-FEB-1995

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 3469, 214-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1060 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Hamicola insolens  
STRAIN: DSM 1800  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 73..924  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 10..72  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 10..924  
US-08-389-423-1

Query Match 37.3%; Score 437.4; DB 2; Length 1060;  
Best Local Similarity 70.4%; Pred. No. 1.3e-78;  
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

55 CCAGAGTGCCTCTTACCTCCGCTTCTGCAACAACCTGCGCTGCACTTCTCTGCTG 114  
5 CCAGATGCGTCTCTCCCTCCCTCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 64  
115 CTTCCGCGGCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 174  
65 CCGTCCGCTGATGAGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 118  
175 GCGGCTGCGCGGAGAGCGCGCTGAGCGCAACCGGCTGAGCGCTGAGCGCTGAGCGCTG 234  
119 GCGGCTGCGCGGAGAGAGGCTCCCTGAGCAACGCTGCTCTTCTGAGCAACGCTGCTC 178  
235 AGCGCTGCTGCACTTCATATGCAATGCGGCTGCA---GCGCGCTGCGCTGCTCT 291  
179 AGCGTATGAGCACTTCAGCGCAAGTCCGCTGCGAGCGCGGCTGCTGCTGCTGCTG 238  
292 GCGCGAGCAAGACTCCCTGCGGCTGAGCAAGCAATGCTGCTGCTGCTGCTGCTGCTG 351  
239 GCGCGAGCAAGACCCCATGAGCTGAGCAAGCAAGCACTGCGCTGCTGCTGCTGCTG 298  
352 GCATGCGCGGCGGATCCGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 411  
299 CATATGCGCGGAGCAATGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 358  
412 CCGGTCGCTGCGCGGAGCAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 471  
359 CCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 418  
472 GAAGTACAGTTCGATGCAATGCGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTG 531  
419 GAGGCAACCACTTGATCTCAACATCCCGCGGCGGCGGCTGCTGCTGCTGCTGCTG 478  
532 GCTGCAAGTTGCGCGGCTCCCGCGGCTCAATATGCGGCGGCTGCTGCTGCTGCTG 591  
479 CTCCCAAGTTGCGCGGCTGCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 538  
592 GCGATCTCTTCCCGCGGCTGCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651  
539 GCGATCGATTCGCGGCGGCTGCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598

652 AGCGGCAACCCGACGTTCAACGTTCCAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 711  
599 AGCGGCAACCCGACGTTCAACGTTCCAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 658  
712 GCTCCGCTGCAAGGCAAGCAAGTTCAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTG 771  
659 GCAACGATGCGCGGAGCAAGCAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 718  
772 GCAACG---TGCAACGCGGAGCGCCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 813  
719 GCAACGCTCTCGGTTCAACGCTTACAGGACCAAGCAAGTGCAGTGCAGTGCAGTGC 778  
814 CGGCGCAAGC---GCTCCGCGCGCGGCGGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 870  
779 GAGCGCGGCAAGTCCAGGCTTACAGCTCCAGGCGGCTGCACTGCTGAGAGTGGCTCAGT 838  
871 GCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 930  
839 GCGGCGCAAGTGCAGGCGGCTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 898  
931 TGAACGACTACTGCGAGTGCCTTAAC 961  
899 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 929

RESULT 12  
US-08-816-915-7  
Sequence 7, Application US/08816915  
Patent No. 6060305  
GENERAL INFORMATION:  
APPLICANT: Royer, John C  
APPLICANT: Moyer, Donna L  
APPLICANT: Yoder, Wendy T  
APPLICANT: Shuster, Jeffrey R  
TITLE OF INVENTION: NON-TOXIC, NON-TOXIGENIC, NON-PATHOGENIC  
TITLE OF INVENTION: FUSARIUM EXPRESSION SYSTEM  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESS: No. 6060305 No. 6060305disk of No. 6060305th America, Inc.  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/816,915  
FILING DATE: 13-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Aprils Dr. Cheryl H  
REGISTRATION NUMBER: 34,086  
REFERENCE/DOCKET NUMBER: 4216, 240-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1060 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 10..924  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 73..924

FEATURE:  
NAME/KEY: sig peptide  
LOCATION: 10..72  
US-08-816-915-7

Query Match 37.3%; Score 437.4; DB 3; Length 1060;  
Best Local Similarity 70.4%; Pred. No. 1.3e-78;  
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 55 CCAAGATGCGCTCTCCGCTTCTGCAACCCCTGCGCTGCACTTCTCTGTG 114  
DB 5 CCAAGATGCGCTCTCCGCTTCTGCAACCCCTGCGCTGCACTTCTCTGTG 64  
QY 115 CCG 174  
DB 65 CCG 118  
QY 175 GCGGCTTGGCG 234  
DB 119 GCGGCTTGGCG 178  
QY 235 AGCGCTTGGCG 291  
DB 179 AGCGCTTGGCG 238  
QY 292 GCG 351  
DB 239 GCG 298  
QY 352 GCG 411  
DB 299 GCG 358  
QY 412 CCGGCTTGGCG 471  
DB 359 CCGGCTTGGCG 418  
QY 472 GAAGTACCG 531  
DB 419 GAAGTACCG 478  
QY 532 GCG 591  
DB 479 GCG 538  
QY 592 GCGCGCTTGGCG 651  
DB 539 GCGCGCTTGGCG 598  
QY 652 AGCG 711  
DB 599 AGCG 658  
QY 712 GCGCGCTTGGCG 771  
DB 659 GCGCGCTTGGCG 718  
QY 772 GCG 813  
DB 719 GCGCGCTTGGCG 778  
QY 814 CCG 870  
DB 779 CCG 838  
QY 871 GCGCGCTTGGCG 930  
DB 839 GCGCGCTTGGCG 898  
QY 931 TGAAGTACCG 961  
DB 899 TGAAGTACCG 929

## RESULT 13

US-09-189-060B-55

Sequence 55, Application US/09189060B

Patent No. 6270968

GENERAL INFORMATION:

APPLICANT: Dalboge, Henrik

APPLICANT: Sandal, Thomas

APPLICANT: Kaupinen, Markus

APPLICANT: Borge, Didrichsen

TITLE OF INVENTION: Method of Providing No. 6270968e1 DNA Sequences

FILE REFERENCE: 4772.204-US

CURRENT APPLICATION NUMBER: US/09/189, 060B

CURRENT FILING DATE: 1998-11-10

PRIOR APPLICATION NUMBER: PCT/DK97/00216

PRIOR FILING DATE: 1997-05-12

NUMBER OF SEQ ID NOS: 74

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 55

LENGTH: 1060

TYPE: DNA

ORGANISM: Humicola insolens

FEATURE:

NAME/KEY: CDS

NAME/KEY: mat. peptide

LOCATION: (73) ... (927)

NAME/KEY: sig. peptide

LOCATION: (10) ... (72)

US-09-189-060B-55

Query Match 37.3%; Score 437.4; DB 3; Length 1060;

Best Local Similarity 70.4%; Pred. No. 1.3e-78;

Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 55 CCAAGATGCGCTCTCCGCTTCTGCAACCCCTGCGCTGCACTTCTCTGTG 114  
DB 5 CCAAGATGCGCTCTCCGCTTCTGCAACCCCTGCGCTGCACTTCTCTGTG 64  
QY 115 CCG 174  
DB 65 CCG 118  
QY 175 GCGGCTTGGCG 234  
DB 119 GCGGCTTGGCG 178  
QY 235 AGCGCTTGGCG 291  
DB 179 AGCGCTTGGCG 238  
QY 292 GCG 351  
DB 239 GCG 298  
QY 352 GCGCGCTTGGCG 411  
DB 299 GCGCGCTTGGCG 358  
QY 412 CCGGCTTGGCG 471  
DB 359 CCGGCTTGGCG 418  
QY 472 GAAGTACCG 531  
DB 419 GAAGTACCG 478  
QY 532 GCGCGCTTGGCG 591  
DB 479 GCGCGCTTGGCG 538  
QY 592 GCGCGCTTGGCG 651

Db 539 GCGATCGTTCCCGACGCTCTCAAGCCGCGCTGCTACTGGGCTTTCAGCTGTTCAAGA 598  
Qy 652 AGCGGACAAACCCGAGCTTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 711  
Db 599 AGCGGACAAACCCGAGCTTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658  
Qy 722 GCTCGGCTGCAAGGAG 771  
Db 659 GCGATCGATCCCGGACAG 718  
Qy 772 GCAACGG-----TGACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 813  
Db 719 GACACGAGCTTCGCGGACAG 778  
Qy 814 GCGGCGAGAG--GTCCTCCGAG 870  
Db 779 GAGGCGGAG 838  
Qy 871 GCGGAG 930  
Db 839 GCGGAG 898  
Qy 931 TGAAG 961  
Db 899 TTAATGAG 929

## RESULT 14

US-09-230-665-1  
Sequence 1, Application US/09230665  
Patent No. 6322595  
GENERAL INFORMATION:  
APPLICANT: Boyer, Stanton L  
TITLE OF INVENTION: Detergent Composition Comprising Two Cellulase  
FILE OF INVENTION: Components, with and without a Cellulose-Binding Domain  
Patent No. 6322595  
CURRENT APPLICATION NUMBER: US/09/230,665  
CURRENT FILING DATE: 1999-01-29  
EARLIER APPLICATION NUMBER: 60/023,125  
EARLIER FILING DATE: 1996-07-30  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 1060  
TYPE: DNA  
ORGANISM: Humicola insolens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (10)..(924)  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: (73)..(924)  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: (10)..(72)  
US-09-230-665-1

Query Match 37.3%; Score 437.4; DB 4; Length 1060;  
Best Local Similarity 70.4%; Pred. No. 1,3e-78;  
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

Qy 55 CGAGATGCGCTCTACTCCCGCTCTGCAACACCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAG 114  
Db 5 CGAGATGCGCTCTACTCCCGCTCTGCAACACCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAG 64  
Qy 115 CCTCGGAG 174  
Db 65 CCTCGGAG 118  
Qy 175 GCGCTGAG 234  
Db 119 GCGCTGAG 178

Qy 235 AGGCTGTCGAGCTTCAATGTCAGTGGAGTGA--CGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 291  
Db 179 AGGCTGTCGAGCTTCAATGTCAGTGGAGTGA--CGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 238  
Qy 292 GCGGAG 351  
Db 239 GCGGAG 298  
Qy 352 GCATCGCGAG 411  
Db 299 CATATCGGAG 358  
Qy 412 CCGGTCGAG 471  
Db 359 CCGGTCGAG 418  
Qy 472 GAAGTACAG 531  
Db 419 GAG 478  
Qy 532 GCTCGAG 591  
Db 479 CCGGTCGAG 538  
Qy 592 GCGATTCCTGCGAG 651  
Db 539 GCGATTCCTGCGAG 598  
Qy 652 AGCGGACAAACCCGAGCTTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 711  
Db 599 AGCGGACAAACCCGAGCTTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658  
Qy 712 GCTCGGCTGCAAGGAG 771  
Db 659 GCGGAG 718  
Qy 772 GCAACGG-----TGACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 813  
Db 719 GACACGAGCTTCGCGGACAG 778  
Qy 814 GCGGCGAGAG--GTCCTCCGAG 870  
Db 779 GAGGCGGAG 838  
Qy 871 GCGGAG 930  
Db 839 GCGGAG 898  
Qy 931 TGAAG 961  
Db 899 TTAATGAG 929

## RESULT 15

US-09-189-028-1  
Sequence 1, Application US/09189028  
Patent No. 6423524  
GENERAL INFORMATION:  
APPLICANT: Raemussen, Gretle  
APPLICANT: Mikelberg, Jan Moller  
APPLICANT: Schuelein, Martin  
APPLICANT: Parker, Shankant A.  
APPLICANT: Hagen, Fred  
TITLE OF INVENTION: A Cellulase Preparation Comprising an  
NUMBER OF SEQUENCES: 33  
TITLE OF INVENTION: Endoglucanase Enzyme  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: No. 6423524C No. 6423524disk of No. 6423524th America, Inc.  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America

ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/09/189,028  
APPLICATION NUMBER: US/09/189,028  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/389,423  
FILING DATE: 14-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 3469.214-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1060 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHEICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: *Humicola insolens*  
STRAIN: DSM 1800  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 73..924  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 10..72  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 10..924  
US-09-189-028-1

Query Match 37.3%; Score 437.4; DB 4; Length 1060;  
Best Local Similarity 70.4%; Pred. No. 1.3e-78;  
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 55 CAGAGTGGCTCTACTCCCGTTCCTTCGACAACTGGCCGCTGCACTTCCTCTGTGTG 114  
DB 5 CCAAGATGGCTTCCTCCCTCTCTCCGTCGCGCGTGTGGCGCCCTGCGGCTGTGG 64

QY 115 CTTCCGCGGCGAGTGGCACTGGCCAGTCCAGAGTACTGGAATCTGTGCAAGCGCTCGT 174  
DB 65 CCTCTCCGCTGATGTGGAG-----GTCCACCCGCTACTGGGACTGTGCAAGCTTCGT 118

QY 175 GCGCTTGGCCCGGAGAGCGCCGCTGAGCCAAACCGCTACAGCGGTGGAGTCCCAACTCC 234  
DB 119 GCGGCTGGGCTCAAGAGAGCTCCCGTGAACCAAGCTGTCTTTCTTCAACGCAACTTCC 178

QY 235 AGGCGCTGTCCGACTTCATATGTCCAGTGGGCTGCAA---CGGCGGCTCGGCTACTTCT 291  
DB 179 AGGCTATCAAGGACTTCGAGCGCAAGTCCGGCTGCGAGCGCGGCGGTGTGCGCTACTCGT 238

QY 292 GCGCGCACTCAACTCTCTGGGCGGTGAACGAACTTCGCTACAGGCTTGGCGCGAGCA 351  
DB 239 GCGCGCACTCAAGCCCAAGGAGTGTGAACGAACTTCGCGGCTCGGCTTGTGCGCACT 298

QY 352 GCATCGCGCGCGGCTCGAACTCTGTGTGTGTGCGGCTGCTACAGGCTCACTTCACTT 411  
DB 299 CTATTCGCGGAGCAATGAGCGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 358

QY 412 CCGGTCCTGTGCGCGGAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 471  
DB 359 CCGGTCCTGT 418

QY 472 GAATTAACCACTTCATATGCAATGCGCCGCGCGCGGTGAGATCTTCAAGCGCTGCA 531  
DB 419 GAGAGCAACCTTCATCAATCAATCCCGCGCGCGGTGAGATCTTCAAGCGATGCA 478

QY 532 GCTCCAGTTCG 591  
DB 479 CTCCCAAGTTCG 538

QY 592 GCGATTCCTTCCCG 651  
DB 539 GCGATTCCTTCCCG 598

QY 652 AGCGCAACCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 711  
DB 599 AGCGCAACCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 658

QY 712 GCTCCGCGTGAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCAAGCGCAAG 771  
DB 659 GCGCGAGTTCG 718

QY 772 GCACGCG-----TGGCAACCGGAGCGCGCAAGCTTCAGCTTCAGCTTCAGCTTC 813  
DB 719 GCAACAGCTTCG 778

QY 814 CGGCGCAAGC---GTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 870  
DB 779 CGAGCG 838

QY 871 GCGGTGGATGGGCTTCAGCGAGTGAACCAACTGTGTGTGTGTGTGTGTGTGTGTGTGT 930  
DB 839 GCGGCGGAGT 898

QY 931 TGAACGACTACTGCGAGTGCCTTAAAC 961  
DB 899 TTAATGACTGTACATGAGTGCCTGTAGAC 929

Search completed: July 7, 2004, 11:41:08  
Job time: 123 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 10:19:37 ; Search time 4330 Seconds  
(without alignments)  
1307.995 Million cell updates/sec

Title: US-10-007-521-11

Perfect score: 1174

Sequence: 1 GAGCAGCAGCCCTCAAGCTG.....AAAAAAAAAAAAAAAAAAAA 1174

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 3163042 seqs, 2412103800 residues

6326084

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PTCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PTCT\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1174	100.0	1174	US-10-007-521-11 Sequence 11, Appl
2	453.6	38.6	1261	US-09-261-329-23 Sequence 23, Appl
3	437.4	37.3	1060	US-09-735-787-1 Sequence 1, Appl
4	437.4	37.3	1060	US-10-138-870-1 Sequence 1, Appl
5	377.2	32.1	890	US-10-369-493-27891 Sequence 27891, A
6	377.2	32.1	1257	US-10-416-328-7 Sequence 7, Appl
7	377.2	32.1	1257	US-10-432-290-18 Sequence 18, Appl
8	369.8	31.5	1154	US-10-007-521-7 Sequence 7, Appl
9	353.6	30.1	936	US-08-841-436A-30 Sequence 30, Appl
10	337	28.7	894	US-10-007-521-3 Sequence 3, Appl
11	334.8	28.5	927	US-10-007-521-5 Sequence 5, Appl
12	315.2	26.8	1473	US-09-735-787-3 Sequence 3, Appl
13	315.2	26.8	1473	US-10-138-870-3 Sequence 3, Appl
14	313.6	26.7	885	US-10-007-521-23 Sequence 23, Appl

15	302	25.7	1132	US-10-007-521-21	Sequence 21, Appl
16	289	24.6	672	US-10-432-290-39	Sequence 39, Appl
17	288.2	24.5	1423	US-10-007-521-9	Sequence 9, Appl
18	274.4	23.4	960	US-10-007-521-1	Sequence 1, Appl
19	260.2	22.2	1043	US-10-432-290-13	Sequence 13, Appl
20	254.6	21.7	913	US-10-007-521-13	Sequence 13, Appl
21	228	19.4	1041	US-10-432-290-12	Sequence 12, Appl
22	213.8	18.2	425	US-10-007-521-25	Sequence 25, Appl
23	210.8	18.0	1031	US-10-007-521-19	Sequence 19, Appl
24	210.8	18.0	1048	US-10-007-521-17	Sequence 17, Appl
25	184	15.7	1017	US-10-432-290-8	Sequence 8, Appl
26	181	15.4	1083	US-10-432-290-10	Sequence 10, Appl
27	158.8	13.5	1101	US-10-432-290-6	Sequence 6, Appl
28	153.6	13.1	1017	US-10-432-290-4	Sequence 4, Appl
29	150.2	12.8	808	US-10-007-521-15	Sequence 15, Appl
30	109.8	9.4	177	US-10-007-521-37	Sequence 37, Appl
31	98.6	8.4	153	US-10-007-521-39	Sequence 39, Appl
32	91	7.8	180	US-10-007-521-41	Sequence 41, Appl
33	85.2	7.3	171	US-10-007-521-51	Sequence 51, Appl
34	84.2	7.2	225	US-10-007-521-31	Sequence 31, Appl
35	83.2	7.1	887	US-08-841-636A-36	Sequence 36, Appl
36	80.8	6.9	165	US-10-007-521-77	Sequence 77, Appl
37	80.8	6.9	159	US-10-007-521-71	Sequence 71, Appl
38	70.8	6.0	147	US-10-007-521-57	Sequence 57, Appl
39	69.8	5.9	177	US-10-007-521-33	Sequence 33, Appl
40	69.2	5.9	413	US-10-437-963-86254	Sequence 86254, A
41	66.2	5.6	608	US-10-424-599-75014	Sequence 75014, A
42	66.2	5.6	419	US-10-437-963-20567	Sequence 20567, A
43	66	5.6	682	US-09-764-847-20	Sequence 20, Appl
44	66	5.6	682	US-10-092-154-20	Sequence 20, Appl
45	66	5.6	682	US-10-092-154-20	Sequence 20, Appl

#### ALIGNMENTS

RESULT 1  
US-10-007-521-11  
; Sequence 11, Application US/10007521  
; Publicatior No. US20030054539A1  
; GENERAL INFORMATION:  
APPLICANT: Schuelein, Martin  
Andersen, Lene N.  
Lassen, Soren F.  
Kauppinen, Markus S.  
Lange, Lene  
Nielsen, Rudy I.  
Ihara, Michiko  
Takagi, Shinobu  
TITLE OF INVENTION: No. US20030054539A1e1 Endoglucanases  
NUMBER OF SEQUENCES: 109  
CORRESPONDENCE ADDRESS:  
ADDRESSER: No. US20030054539A10 No. US20030054539A1disk of No. US200300:  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/007521  
FILING DATE: 10-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/651,136  
FILING DATE: 21-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728



REFERENCE/DOCKET NUMBER: 4366-200-US  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1174 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: cDNA  
FEATURE:

NAME/KEY: CDS  
LOCATION: 60..956

SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-10-007-521-11

Query Match 100.0%; Score 1174; DB 15; Length 1174;  
Best Local Similarity 100.0%; Pred. No. 1e-265;  
Matches 1174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GAGCAGACCCCTCAAGCTGTACAGTTTCCACCCCGCTCTCTTTCTTCCGCCCCCAAGA 60
DB 1 GAGCAGACCCCTCAAGCTGTACAGTTTCCACCCCGCTCTCTTTCTTCCGCCCCCAAGA 60
QY 61 TCCGCTCTACTCCCGTTCTTCCACAACTTCCGCTCTCTCTCTCTCTCTCTCTCTCT 120
DB 61 TCCGCTCTACTCCCGTTCTTCCACAACTTCCGCTCTCTCTCTCTCTCTCTCTCTCT 120
QY 121 CGGCCAGTGGCAGTGGCAGTGCACAGATACCTGCGGACCGCTGCGTGGCTT 180
DB 121 CGGCCAGTGGCAGTGGCAGTGCACAGATACCTGCGGACCGCTGCGTGGCTT 180
QY 181 GGCCTCGGAAAGCCCGCTGACGCAACCGGCTTACGCTGCGATGCGCACTTCCAGC 240
DB 181 GGCCTCGGAAAGCCCGCTGACGCAACCGGCTTACGCTGCGATGCGCACTTCCAGC 240
QY 241 TGTCCGACTTCAATGTCCAGTGGGCTGGCAAGCGGCGCTTCCGCTTCTCTGCGGACC 300
DB 241 TGTCCGACTTCAATGTCCAGTGGGCTGGCAAGCGGCGCTTCCGCTTCTCTGCGGACC 300
QY 301 AGACTCCCTGGGCGGTGAAACCAATCTCCCTCAACCGCTTCCGCGCGACGAGCATCG 360
DB 301 AGACTCCCTGGGCGGTGAAACCAATCTCCCTCAACCGCTTCCGCGCGACGAGCATCG 360
QY 361 GCGGCTCGAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
DB 361 GCGGCTCGAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 421 TCGCGCGAAGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
DB 421 TCGCGCGAAGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
QY 481 AGTTGATATGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 AGTTGATATGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 TCGGCGGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 TCGGCGGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 TCCCGCGCGCGCTCAAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
DB 601 TCCCGCGCGCGCTCAAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
QY 661 ACCCGAGCTTCAAGCTTCAAGAGGTGAGTCCCGCGCGCGCGCGCGCGCGCGCGCGCT 720
DB 661 ACCCGAGCTTCAAGCTTCAAGAGGTGAGTCCCGCGCGCGCGCGCGCGCGCGCGCGCT 720
QY 721 GCAAGCGCAAGACATCCAGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
DB 721 GCAAGCGCAAGACATCCAGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
```

```
QY 781 GCACCGGAGACGCTCCAGCTGCACTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCA 840
DB 781 GCACCGGAGACGCTCCAGCTGCACTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCA 840
QY 841 GTGGCTGCAAGCTTCAAGAGGTGAGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 GTGGCTGCAAGCTTCAAGAGGTGAGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 CTTGTGTCTTCCGACCACTTCCGACAGATGAAAGACTACTGCGAGTGGCTTAA 960
DB 901 CTTGTGTCTTCCGACCACTTCCGACAGATGAAAGACTACTGCGAGTGGCTTAA 960
QY 961 CAGCTTTTCCAGAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1020
DB 961 CAGCTTTTCCAGAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1020
QY 1021 TTTTGAAGCGCTCAATACATACATACATACATACATACATACATACATACATACAT 1080
DB 1021 TTTTGAAGCGCTCAATACATACATACATACATACATACATACATACATACATACAT 1080
QY 1081 TCTCAGACCGACTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1140
DB 1081 TCTCAGACCGACTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1140
QY 1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174
DB 1141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1174
```

#### RESULT 2

US-09-261-329-23  
Sequence 23, Application US/09261329  
Publication No. US2003092097A1  
GENERAL INFORMATION:  
APPLICANT: Andersen, Kim  
APPLICANT: Schuelein, Martin  
APPLICANT: Christiansen, Lars  
APPLICANT: Damgaard, Bo  
APPLICANT: Von der Osten, Claus  
TITLE OF INVENTION: Cellulase Variants  
FILE REFERENCE: 4887,204-US  
CURRENT APPLICATION NUMBER: US/09/261,329  
CURRENT FILING DATE: 1999-03-03  
EARLIER APPLICATION NUMBER: 1013/96  
EARLIER FILING DATE: 1996-09-17  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 23  
LENGTH: 1261  
TYPE: DNA  
ORGANISM: Humicola grisea

Query Match 38.6%; Score 453.6; DB 10; Length 1261;  
Best Local Similarity 74.9%; Pred. No. 1.5e-96;  
Matches 583; Conservative 0; Mismatches 189; Indels 6; Gaps 1;

```
QY 26 TTTTCAACCCCGCTCTTTTCTTGTGAGCCCGAGAGTGGCTTACCTCCGTTCTTGAC 85
DB 18 TTTTATTCACGCTCATTTATTTCAAACTCAATATGCGCTGCTCTATTTTCGAC 77
QY 86 AACCTGGGCGGCTGACCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 145
DB 78 GGCCTGGGCGGCTGCTTCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 131
QY 146 GAGATCTTGGGACTGTGCAAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 205
DB 132 CAATATCTGGGACTGTGCAAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 191
QY 206 ACGGCTTCAAGCGGAGTGCATCTTCAAGCGGCTGTGTGTGTGTGTGTGTGTGTGT 265
DB 192 GCGTGTCTTCAAGCGGAGTGCATCTTCAAGCGGCTGTGTGTGTGTGTGTGTGTGT 251
```

QY 286 TTGAAAGAGGCTGAGCTACCTCCGAGCGCAACAATCTCCCTGAGGAGTGAAGACA 325  
 |||||  
 Db 252 CTGAGATGAGGCTGAGCTCTTTTCGTGTGCTGACAGACCCCTTGAGCTCTAAGAGA 311  
 QY 356 TCTGAGCTTAAGGCTTCGCGCGACAGAGCATGCGCGAGGTCCAAATCTCGTGTGCTG 385  
 |||||  
 Db 312 TGTGAGCTTAAGGCTTCGCTGACAGGCTATTGTTGGAGGTGATCGAAGCTCTGTGCTG 371  
 QY 446 GTCAACAGACACTGAGCGGCACTGAGGAAGTAACTGTTGATATGCGCATGCGCGAGG 505  
 |||||  
 Db 432 GTCAACCTAAACCGGCGGCGCATCTCGGAGAGCAACATTTGACCTCGAGTTTCAGGCGG 491  
 QY 556 CGGCGTGGGATCTTCAACCGGCTGAGCTCGAGTTGAGGAGCTTCCCGGAGCTCAATA 565  
 |||||  
 Db 492 CGGTTGCGGCACTTTGATATGAGTGAACCCCGCGAGTTGAGAGGCTCTGCTGAGAAAGCTA 551  
 QY 566 CGGAGGATTTGTTGTCGAGGACAGAGGAGATTCCTTCCCGGCGCGCTCAAGCGGCTG 625  
 |||||  
 Db 552 CGATGAGATCTAAGACCGGCGGCGATCTCGGAGAGCAACATTTGACCTCGAGTTTCAGGCGG 611  
 QY 626 CCAATGAGCGGTTTGAATGCTTCGAGAGCGCGCAACCCGAGTTTCAAGTTCAAGCAAGT 685  
 |||||  
 Db 612 CATTGCGGCTTCAATTTGATTGATTCAAGAGCGCGAACAACCGACCTTTACCTTCAACAAGT 671  
 QY 686 GCAATGCCCCCGGAGAGATCGTTGCCGCTCGGCTGCAAGCGCAACAAGATCTCAAGCTT 745  
 |||||  
 Db 672 GCAATGCCCCCGGAGAGCTTTGTTGCCAGAGACCGGCTGCAAGCGCAAGAGATGAGGCAACT 731  
 QY 746 CCGCGCTTCAACCCCGCAAGCGTGTGAGCAAGGTGCAACCGGAGAGCCCAAGCTGACT 803  
 |||||  
 Db 732 CCGCGCTTCAACGCCCCCGGAGGTGAGAACACCGGAGGTGAGTCAAGTCAAGTCAACT 789  
 |||||  
 RESULT 3  
 US-09-735-787-1  
 ; Sequence 1, Application US/09735787  
 ; Patent No. US20010036910A1  
 GENERAL INFORMATION:  
 APPLICANT: Rasmussen, Grethe  
 Mikkelsen, Jan Moller  
 Schueleid, Martin  
 Packar, Shankant A.  
 Hagen, Fred  
 TITLE OF INVENTION: A Cellulase Preparation Comprising an  
 Endoglucanase Enzyme  
 NUMBER OF SEQUENCES: 33  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: No. US20010036910A1O No. US20010036910A1disk of No. US20010036910A1  
 STREET: 405 Lexington Avenue, 64th Floor  
 CITY: New York  
 STATE: New York  
 COUNTRY: United States of America  
 ZIP: 10174-6402  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/735,787  
 FILING DATE: 13-Dec-2000  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/189,028  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Landitis, Elias J.  
 REGISTRATION NUMBER: 33,728  
 REFERENCE/DOCKET NUMBER: 3469, 214-US

	TELECOMMUNICATION INFORMATION:
?	TELEPHONE: 212-867-0123
?	TELEFAX: 212-878-9655
?	INFORMATION FOR SEQ ID NO: 1:
?	SEQUENCE CHARACTERISTICS:
?	LENGTH: Humicola insolens
?	TYPE: nucleic acid
?	STRANDEDNESS: single
?	TOPOLOGY: linear
?	MOLECULE TYPE: cDNA
?	HYPOTHETICAL: NO
?	ORIGINAL SOURCE:
?	STRAIN: DSM 1800
?	FEATURE:
?	NAME/KEY: mat_peptide
?	LOCATION: 73..924
?	FEATURE:
?	NAME/KEY: sig_peptide
?	LOCATION: 10..72
?	FEATURE:
?	NAME/KEY: CDS
?	LOCATION: 10..924
?	SEQUENCE DESCRIPTION: SEQ ID NO: 1:
?	US-09-735-787-1
Query Match	37.3%; Score 437.4; DB 9; Length 1060;
Best Local Similarity	70.4%; Pred. No. 9,2e-93;
Matches 653; Conservative 0; Mismatches 246; Indels 30; Gaps 4	
Dy	55 CGAGATGGGCTCACTCCCGTTCGTGCACAAACCCTGAGCGTGACATCCTCTGATCG 114
Dd	5 CCAAGATGGTCTCTCCCCCTCCCTCCCGTCCGCCGTTGAGCGCCCTGCGGATGTGG 64
Dy	115 CTTCGCGGCGCATGGGAGTAGGACATCCACAGATACTGGAACTGTGCAGAGCGTGT 174
Dd	65 CCTTCGCGCGCATGGAGGAG-----GTCCACCGGCTACTGGAGCTGGAAGCTTTCG 118
Dy	175 GCGCTTGGCCGCGAAGAAGGCGCGCTGCACCAACCGGTCTACGCTGGAGTAACAATTCC 234
Dd	119 GGCGCTGGCCAMAAAGGCTCCCGTGAACCAAGCTGTCTTTTCTGCACACGCAACTTCC 178
Dy	235 AGCGCTGTCCGACTTCAATGTCCAATGTCCAGTCCGGCTTGCAA---CGCGGCTCCGCTTACTCT 291
Dd	179 AAGTATACAGGAATTTCAGACGAAGTCCGGCTGGAACCGGCGGATGTGCTCACTCT 238
Dy	292 GCCCGCACCAACTCTCCCTGGGCGGTGAACGACAAATCTGCGCTTACGSGCTTCCGCGACGA 351
Dd	239 GCACCGACCAACCCCACATGGGCTGTGGAACGACAACTTCCGCTCGGTTTGTGCAACT 298
Dy	352 GCATGCGCGGCGGAGTCCGAATCTTCGAGTGTGCGGCGCTGTACGCGCTCAACTCACT 411
Dd	299 CTATTTCGCGGACGAATGAGGCGGAGCTGATGTGGCGCTGTCTACAGCTCACTTCAACT 358
Dy	412 CCGGTCCCGTGCACCGGCAAGCAATGTGTGATGTCACTTCAAGCACTGGACCGGCACTTGG 471
Dd	359 CCGGCTGTGTTGCTGGGAAGAAATGTGTGTCACTCAACACACTGGCGGATCTTG 418
Dy	472 GAAGTACCAATGTGATATGCGCATGTCCGCGGCGGCGGTGGGCACTTTCAAGGCTGCA 531
Dd	419 GCAGCAACCACTTCATCTCAACATTCGCCGCGGCGGAGTGTGCACTTTCACAGATGCA 478
Dy	532 GCTGCGAGTGGCGGCGCTCCCGCGGCGTCAATAAGGCGCAATTCGTGTGCGGACAGT 591
Dd	479 CTCGCCAGTTGGAGGATCTGAGCCCGGACAGGCTTACGCGGCACTCTGTCGCCGAAGAT 538
Dy	592 GCGATTCCTTCCCGGCGGCTCAAGCCCGGCTGCAGTGGGATTTGACTGGTTCAGA 651
Dd	539 GCGATGGTTCGCCAGAGGCCCTCAAGCCCGGCTGTCACTGGCGCTTCACTGATGATCAGA 598
Dy	652 AACCGCAACACCGAGTTCAAGTTCACAGAGTGGAGTGGACCCCGGAGTGTGTTGCC 711
Dd	599 AACCGCAACATCGACTTCAAGTTCCTTCGACAGTTCACAGTCCCAAGCCGAGTGTGCTG 658

QY 712 GCTCCGCTGCAAGCGCAACGACCTCCGCTTCCGCTTCAACCCCAAGGGTG 77-  
| | | | |  
Db 653 GACCCGAGTCCGCCGCAACGACGGAATTCTCCCTGCGTCAAGTCCCTCCAGCA 718  
| | | | |  
QY 772 GCACGCG-----TGGCAACGCGGAGCGCCGACCTGACCTGCGCTGGGT 813  
| | | | |  
Db 719 GCACGAGCTCCGCGTCAACGACCTTACGAGACGACCAACGCTTCACTCCACACT 778  
| | | | |  
QY 814 CCGGCCAAGC---GTCTCCCGCGCGCGAGTGGCTGCAAGTCTCAAGAGTGGGCTCAGT 870  
| | | | |  
Db 779 CGAGCCCGCCAGTCCAGCTTACGACTCCGCGCTGCACTGCTGAGAGGCTGGGCTCAGT 838  
| | | | |  
QY 871 CGCGTGGCATGCGCTTCAAGGAGTACGACCTTGTCTCTGCGACCACTGCGCAAGT 930  
| | | | |  
Db 839 CGCGCGGAGTGGCTGAGAGCGGCTGCAACCTGCTGCTGCTGCGTGGCACTTGCACAGAA 898  
| | | | |  
QY 931 TGAACGACTACTACTGCGAGTCCCTTAAC 961  
| | | | |  
Db 899 TTAATGACTGGTACATCACTGCTGTAGAC 929  
| | | | |

## RESULT 4

US-10-138-870-1  
; Sequence 1, Application US/10138870  
; Publication No. US20030119167A1

## GENERAL INFORMATION:

APPLICANT: Rasmussen, Grethe  
Mikkelsen, Jan Koller  
Schulein, Martin  
Pakkar, Shankant A.  
Hagen, Fred

TITLE OF INVENTION: A Cellulase Preparation Comprising an  
Endoglucanase Enzyme

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. US20030119167A10 No. US20030119167A1disk of No. US200301191  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York

COUNTRY: United States of America  
ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10138,870

FILING DATE: 03-May-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/735,787

FILING DATE: 13-Dec-2000

APPLICATION NUMBER: 09/189,028

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 3469,214-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1060

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ORIGINAL SOURCE:

STRAIN: DSM 1800

NAME/KEY: mat\_peptide  
LOCATION: 73..924  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 10..72  
NAME/KEY: CDS  
LOCATION: 10..924  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-138-870-1

Query Match 37.3%; Score 437.4; DB 15; Length 1060;  
Best Local Similarity 70.4%; Pred. No. 9,28-93;  
Matches 655; Conservative 0; Mismatches 246; Indels 30; Gaps 4;

QY 55 CCAAGATGGCTCTTACTCCGTTCTTCCGCAACCTCGCGCTGCACTTCTCTGCTG 114  
| | | | |  
Db 5 CCAAGATGGCTCTTCCCGCTTCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 64  
| | | | |  
QY 115 CTTCCGCGCGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 174  
| | | | |  
Db 65 CTTCCGCGCTGATGGAG-----GTCCACCTCTACTGGAGCTGGCAAGCTTGGT 118  
| | | | |  
QY 175 GCGCTTGGCGCGGAGAGCGCGCTGAGCGCAACGCTCTACGCGTGCATGCAACTTCC 234  
| | | | |  
Db 119 GCGGCTGGCGCGCAAGAGGCTCCGCTGACGCGCTGCTTCTTCCGCAACGCAACTTCC 178  
| | | | |  
QY 235 AGGCGCTTGGCTCTTCAATGTCAGTGGCTGCA--CGCGGCTTGGCTTCTTCT 231  
| | | | |  
Db 179 AGGATTCACGAGCTTGGAGCGCAAGTCCGCTGAGCGCGCGGCTGCTTCTTCT 238  
| | | | |  
QY 232 GCGCGGAGGAGCTCCCGCGGCTGAGCAACATCTGCTGCTGCTGCTGCTGCTGCTG 351  
| | | | |  
Db 239 GCGCGGAGGAGCTCCCGCGGCTGAGCAACATCTGCTGCTGCTGCTGCTGCTGCTG 298  
| | | | |  
QY 299 CTATTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 358  
| | | | |  
Db 412 CCGGCTCCCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 471  
| | | | |  
QY 359 CCGGCTCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 418  
| | | | |  
QY 472 GAAGTAAACGATGATATGCGCATGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGG 531  
| | | | |  
Db 419 GAGGAGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 478  
| | | | |  
QY 532 GCTCGAGTGGCGCGGCTTCCCGCGGCTCAATACGCGGCGGCGGCGGCGGCGGCGG 591  
| | | | |  
Db 479 CTCGCCAGTTCGCGCGGCTTCCCGCGGCTCAATACGCGGCGGCGGCGGCGGCGG 538  
| | | | |  
QY 592 GCGATTCCTTCCCGCGGCGGCTTCCCGCGGCTCAATACGCGGCGGCGGCGGCGGCGG 651  
| | | | |  
Db 539 GCGATTCCTTCCCGCGGCGGCTTCCCGCGGCTCAATACGCGGCGGCGGCGGCGGCGG 598  
| | | | |  
QY 652 ACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 711  
| | | | |  
Db 599 ACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 658  
| | | | |  
QY 712 GCTTCGCGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 771  
| | | | |  
Db 659 GCAACGAGATGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 718  
| | | | |  
QY 772 GCACGCG-----TGGCAACGCGGAGCGCCGACGCTGCACTGCGCTGGGT 813  
| | | | |  
Db 719 GCACGAGCTTCCGCTCAACGACCTTACGAGACGACCAACGAGTCCCTCCAGCACT 778  
| | | | |  
QY 814 CCGGCCAAGC---GTCTCCCGCGCGCGAGTGGCTGCAAGTCTCAAGAGTGGGCTCAGT 870  
| | | | |  
Db 779 CGAGCCCGCGAGTCCAGCTTACGACTCCGAGGCGCTGCACTGCTGAGAGGCTGGCTCAGT 838  
| | | | |  
QY 871 GCGGAGGAGTGGCTTCAAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 930  
| | | | |



Db	348	CGCGAACAAGACCCCAATGGGCTGTGAAGAAGACATTCGCGTTTGCTGTTGGCGACCTTC	407
Ox	353	CATGCGCGGAGGTGCGAATCTCTGCTGCTGCGCTGTCTAAG-----	396
Db	408	TATTCGCGGACGACATAGGCGGCGCTGTGCTCGGCTGTACAGAAAGATCTTGTGCG	467
Ox	397	-----GCTCACTTCACTTCGCGT	416
Db	468	GTCGTGAACACTGTGCAAGGAGTAAGCACTTAACCACTTCCCAAGCTCACTTCAATCCGAT	527
Ox	417	CCCGTGCACGAGCAAGACATGTGTGTGACGTCAACAGACACTGCGCGACCTGAGAGT	476
Db	528	CTGTGTGTGCTGCAAGAAAGATGATGTCGATCAAGTCAACAGACTGCGCGGTATCTTGACAG	587
Ox	477	AACCAAGTGTGATATGCGCATGTCGCGCGCGCGGTGTGGGCAATCTTCAAGGACTGACGTGC	536
Db	588	AACCACTGTGATCTTCAACATCCCGGAGCGGCGGTGTGGCAATCTTCAAGAGTAGCACTTCC	647
Ox	537	CAGTTCGAGCGGCTTCCCGGCGGCTCAATACGAGCGGCAATTTGTGTGCGGACCAAGTGCAT	596
Db	648	CAGTTCGAGCGGCTTCCCGGCGGCTCAAGGCGCTAAGGCGGATCTGTCTCCGCAAGAGTGGAT	707
Ox	597	TCTTTCCTCCCGCGCGCTCAAGCTCCGCTGCGCAATGACGTGCTTGACTGTTCAGAACGCC	656
Db	708	CGGTTCCCGCGAGCGCTTCAAGCGCGGCTGTACTGTGCGCTTCACTGTGTTCAAGAACGCC	767
Ox	657	GACAAACCGCAAGTTCAGTTCACAGCAAGTGTCAAGTGTGCGCGGACAGATTCGTTGCCGCTCC	716
Db	768	GACAAACCGCAAGTTCAGTTCCTGTGTAGTGCATATGCGACAGCGAGCTGTGCTGTGCAACC	827
Ox	717	GACTTCGAAGCGAAGCAAGCACTTCAGCTTCCCTGTTCATCCCGCCCAAGCGGTGCGAAC	776
Db	828	GGATGTCCGCGCGCAACGACAGCGGCAACTCCCTGCGCTGTCAAGATCCCTCCAGACGACAC	887
Ox	777	GG-----TGGACAAGGAGAGGCCACAGTGCATGTGCGCTGTGCGT	815
Db	888	AGCTTTCGCGTGTGCGCAGCTTACAGTACAGAGCACTTCACTTCACCACTTCGAGC	947
Ox	816	GAGCGAAGCTTCTCCCGCGCGGCGGAGTGTGTGACGCTCTCAAGAGTGTGGCTTACGTGCGGT	875
Db	948	CCGCGCGCTCAAGCTTACAGATCTCCAGAGGCTGACATGCTGAGAGGTGTGGCTAATGCGAG	1007
Ox	876	GGCATCGGCTTACAGGAGTGTACCACTGTGTCTGTGCGACCACTTCCCAAGTGTGAAC	935
Db	1008	GGCAATGTGCTGTGAGGCTGTGCACCACTGTGTGTGTGTGCGACACTGTGCACAGCAATTAAT	1067
Ox	936	GACTACTACTGTGCAGTGTCTCTTAAC	961
Db	1068	GACTGTGACCACTAATGTCTGTGAAC	1093
RESULT 7			
US-10-432-290-18			
/ Sequence 18, Application US/10432290			
/ Publication No. US20040043400A1			
GENERAL INFORMATION:			
APPLICANT: MEIJI SEIKA KAISEI, LTD.			
TITLE OF INVENTION: Endoglucanase Enzyme from Zygomycetes Having Deletion of			
/ TITLE OF INVENTION: Cellulose Binding Domain			
/ FILE REFERENCE: YK032510 PCT			
/ CURRENT APPLICATION NUMBER: US/10/432,290			
/ CURRENT FILING DATE: 2003-05-20			
/ PRIOR APPLICATION NUMBER: PCT/JP01/10188			
/ PRIOR FILING DATE: 2001-11-21			
/ NUMBER OF SEQ ID NOS: 43			
/ SOFTWARE: PatentIn Ver. 2.0			
/ SEQ ID NO 18			
/ LENGTH: 1257			
/ TYPE: DNA			
/ ORGANISM: Humicola insolens			
/ FEATURE:			
/ NAME/KEY: intron			
/ LOCATION: (453)..(509)			

US-10-432-290-18  
 Query Match 32.1%; Score 377.2; DB 13; Length 1257;  
 Best Local Similarity 66.1%; Ped. No. 1,36-78;  
 Matches 652; Conservative 0; Mismatch 248; Indels 86; Gaps 4.

QY	56	CAGGAGGCGCTCTACTCTCCGGTCTTCCGACAAACCCGCGCGCTGACACTTCTCCGTGGCC	115
DB	114	CAGAGTACGCTTCTCCCTCTCTCTCCGCTCCGCGCTTGGCCGCGCTTGGCGGTGGCC	173
QY	116	CTCCGAGCCAGTGGCAGTGGCCAGTCCACGAGATTACTGGACTGTCGACGCGGTGG	175
DB	174	CTTTGCGCGCTGATGGCA-----AGTCCACCCGCTACTGGAGACTGCTGCAGACCTTGGTG	227
QY	176	CGCTTGGCCCGCGAGAAAGCCGCGCTTCAAGCACTCGGTCTAAGCTTGGAGATGCAACTTCCA	235
DB	228	CGGCTGGGCGCAAGAAAGGCTCCGTTGAACAGCGCTTCTTCTCTGCAAGCGCAACTTCCA	287
QY	236	CGCGCTGTCCGACTTCAATATGTCAGATGGGGCTGCAAA--CGGCGGCTCGGCTACTCTTG	292
DB	288	CGGCTCTCATGACTTGGAGGCCAAGTCCGGCTGGCGAGCGCGGGGCGGTGGCTGCTACTGGTG	347
QY	293	CGCGGACGACGACTCTCCCTGGCGGTGGAACGACAACTCTCGCTTAAGGCTTGGCGCGAGAG	352
DB	348	CGCGGACGACGACCCGAGGCGGTGGAACGACAACTTCCGCTTGGCTTGGCTGCGACCTTC	407
QY	353	CATGGCGGCGGGGTCCGAATCTCGTGGTCTGACGCTGCTAGCG-----	396
DB	408	TATTTGCGGCGAGCAATGAGCGGGGTGGTGGTGGCTGGTGGTAAAGCTTTGGTGGC	467
QY	397	-----CGTCCACTTCACTTCCGGT	416
DB	468	GTGTGTATACACTGTGGCAGGCTAGCACTTAACACTCCGACGCGTCAACTTCAATTCGGT	527
QY	417	CCGCTGCGCGGAGAAATAGGTGGTGGATGAACAGACACTGGCGGACACTGAGAAAT	476
DB	528	CTGTGTGTGGAGAAAGATGGTGGTCCATGACACACACTGGCGGTGAATCTTGGCAGC	587
QY	477	AACGAGTTCGATATGCGCAATGCCCGCGCGGCGGTGGGCAATCTTCAACGCGCTGCAGCTGG	536
DB	588	AACCACTTCGATTCACATATCCCGCGCGGCGGTGGGCAATCTTCCAGCGATGCACTCC	647
QY	537	CAGTGTGGGCGCTCCCGCGCGCTCAATTAAGCGGCGCAATTTGGTGGCGGACGATGGAT	596
DB	648	CAGTGTGGGCGGTCTCCCGCGCGGCTCAATTAAGCGGCGCAATTTGGTCCGACGATGGAT	707
QY	597	TCCCTCCCGCGCGCTCAAGCCCGGCTTGGCAGTGGCGGTGAATGATGTTCCAGAACGCC	656
DB	708	CGGTTCCCGAGGCGCTCAAGCCCGGCTTGGCAGTGGCGGTGAATGATGTTCCAGAACGCC	767
QY	657	GACAACTCGAGCTTCAAGTTCCAGAGGTGAGTCCCCCGCGAGATGGTTGGCCCGGCTCC	716
DB	768	GACAACTCGAGCTTCAAGCTTCCGTAAGTCAAAATGCGACGCGAGGCTGGTGGCTCGACCC	827
QY	717	GGCTGCAAGCGCAACAGCACTCCAGCTTCCCGGTCTTCAACCCCGCGCAAGCGGTGGCAAC	776
DB	828	GGAATGCGCGCGAAGAGAGCGGGAATCTTCCGTGGGTTCAGAAATCCCTCCAGGAGACCC	887
QY	777	GG-----TGGACCGGAGGCGCCAGTGTGACTGGGCTGGGCTGG	815
DB	888	AGCTCTCCGCTGGCGCAGGCTACAGTAACGAGCAACAATCTTCCACCTTCCACCTTCCAGC	947
QY	816	GGCGACAGCTCTCCCGCGCGGCGGAGTGGCTGACGCTTCCAGAAAGTGGGCTCAATGGGCT	875
DB	948	CGGCGCGTCAAGCTTCAAGTCCGACGCGGTGCACTGCTGAAGGTGGGCTCAATGGGCGC	1007
QY	876	GGGATCGGCTTCAAGGATGACGACCTGTGTCTTGGCAACAACCTTGCAGAGATTGAAC	935
DB	1008	GGGAAAGGCTGGAGACGGCTGGCAACAACCTGGGTGGCTGGCAACAACGAAATTAAT	1067
QY	936	GACTACTACTCGCAGTGGCTCTTAAAC	961
DB	1068	GACTGTGATCATCATGATGCTCTTAAAC	1093

## RESULT 8

US-10-007-521-7

Sequence 7, Application US/10007521

Publication No. US20030054539A1

## GENERAL INFORMATION:

APPLICANT: Schlein, Martin

Andersen, Lene N.

Lassen, Soren F.

Kauppinen, Markus S.

Lange, Lene

Nielsen, Ruby I.

Ihara, Michiko

Takagi, Shinobu

TITLE OF INVENTION: No. US20030054539A1 Endoglycanases

NUMBER OF SEQUENCES: 109

CORRESPONDENCE ADDRESS:

ADDRESS: No. US20030054539A10 No. US20030054539A1disk of No. US200300545

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-90S/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/10/007,521

APPLICATION NUMBER DATA:

FILING DATE: 10-Dec-2001

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/651,136

FILING DATE: 21-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4366,200-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1154 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 51..935

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-007-521-7

Query Match 31.5%; Score 369.8; DB 15; Length 1154;

Best Local Similarity 60.6%; Pred. No. 7.1e-77;

Matches 666; Conservative 0; Mismatches 417; Indels 16; Gaps 3;

QY 83 CACAACCTGGCGCTGCACTTCTCTGCTCGCTCGCGGCGCAAGTGGAGTGCACATC 142  
DB 62 CAGCATCTTATGCGGCTGTGTCCGGGCTGTGTGTGCTGCAAGACTCTGGCTGTGCGCATC 121  
QY 143 CACGAGATTAATGAGTCTGCTGCAAGCGCTGTGCTTGGCCCGGAAAGGCGCGCTAG 202  
DB 122 AACGAGTACTGGAGTCTGTGCAAGCGCTGTGCTGCTGAGTGAAGAGCGGCTGAG 181  
QY 203 CGAACGGCTCTAGCGCTGTGCAAGTCTTCCAGCGCTGTGCAAGTCTTCAAGTCCATC 262  
DB 182 CGGCGCGGTCAAACTGTGCAAGTGAAGCAAGCGCGCTTTCGCGCGGCTGTGAGCG 241  
QY 263 GGGCTGCAAGCGCGGCTCGGCTTACTCTGCGCGCAAGCAAGTCTTCTGCGCGGTGAAGA 322

DB 242 CTGGACCCCAAGCGGCTTGTGATTTCACTGTGCAACGACCAAGCTTGGCGTAAACGA 301  
QY 323 CAATTCGCTTACGCTTGTGCTGCGGACGACGACATCGCGCGGCTTCCGAATCTCTGTG 382  
DB 302 CAATGTGCTTACGCTTGTGCGGACGCTTCCCTGTGTGGCAATGAAGCGCTGTG 361  
QY 383 CTGGCGCTGTGAGGCTCACTTCACTTCCGCTCCGCTCGCGCAAGCAATGTG 442  
DB 362 CTGGCGCTGTGAGGCTTCAATTCATTCGCGCGCTTCTGTGCGCAAGCAATGTG 421  
QY 443 GCAATCAAGCAAGTGTGCGGCGGACGCGGAGTACAGTGTGCAATTCGCGCGCG 502  
DB 422 GCAATCAAGCAAGTGTGCGGCGGACGCGGAGTACAGTGTGCAATTCGCGCGCG 481  
QY 503 CGGCGCGGTGCGGCTTCAAGCGCTTCAAGTGTGCGGCGCGCGCGCGCGCGCG 559  
DB 482 TGGAGGCTTGTGCGGCTTCAAGCGCTTCAAGTGTGCGGCGCGCGCGCGCGCG 541  
QY 560 TCAATTAAGCGGCTTGTGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAGCC 619  
DB 542 CGGCTTACGCGGCTTCAAGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAG 601  
QY 620 CGGCTTACGCGGCTTCAAGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAG 679  
DB 602 CGGCTTACGCGGCTTCAAGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAG 661  
QY 680 CGAGTGTGCAAGTGTGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAG 739  
DB 662 CGGCTTCAAGTGTGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAG 721  
QY 740 CAGCTTCCGCTTCAAGTGTGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAG 799  
DB 722 CAGCTTCCGCTTCAAGTGTGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAG 772  
QY 800 GACTGCGCGCTTCAAGTGTGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAG 859  
DB 773 CAGCTTCAAGTGTGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAG 832  
QY 860 GTGGGCTCAAGTGTGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAG 919  
DB 833 CTGGGCTCAAGTGTGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAG 892  
QY 920 CTGGGCTCAAGTGTGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAG 979  
DB 893 CTGGGCTCAAGTGTGCGGCGGACGCGGAGTGTGCTTCCGCGCGCGCTGAG 952  
QY 980 GCGGAGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1039  
DB 953 GCGGAGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1012  
QY 1040 TACATTAACCTTCAATTTTGTGATAG---CAGCGCGGTACATCTCAACCGACTTT 1095  
DB 1013 ATGCGATTTATTAATTTTGTGATAG---CAGCGCGGTACATCTCAACCGACTTT 1072  
QY 1096 GGGGCGGAGTGTGCGGCGGACGCGGAGTGTGCGGCGGAGTGTGCGGCGGAGTGTG 1155  
DB 1073 GACCTAGCGAGTGTGCGGCGGACGCGGAGTGTGCGGCGGAGTGTGCGGCGGAGT 1132  
QY 1156 AAAAAAAAAAAAAAAAAA 1174  
DB 1133 AAAAAAAAAAAAAAAAAA 1151

RESULT 9  
US-08-841-636A-30  
Sequence 30, Application US/08841636A  
Publication No. US20020168751A1  
GENERAL INFORMATION:  
APPLICANT: Miettinen-Ononen, Arja  
APPLICANT: Lomdasborough, John  
APPLICANT: Venmaaper, Jari  
APPLICANT: Haakana, Hei

APPLICANT: M ncy1, Arja  
APPLICANT: Lantto, Raija  
APPLICANT: Elvoinio, Minna  
APPLICANT: Joutsinki, Vesa  
APPLICANT: Paloheimo, Marja  
APPLICANT: Suominen, Pirkko  
TITLE OF INVENTION: NOVEL CELLULASES, THE GENES ENCODING THEM AND  
TITLE OF INVENTION: USBS THEREOF  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/841,636A  
FILING DATE: 30-Apr-1997  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 60/305,335  
FILING DATE: 17-Oct-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 60/007,926  
FILING DATE: 04-Dec-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 60/020,840  
FILING DATE: 28-JUN-1996  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/732,181  
FILING DATE: 16-Oct-1996  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/Fin6/00550  
FILING DATE: 17-Oct-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Timothy J. Shea, Jr.  
REGISTRATION NUMBER: 41,306  
REFERENCE/DOCKET NUMBER: 1716.0510005/MAC/LUS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)371-2600  
TELEFAX: (202)371-2540  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 936 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Melanocarpus albomyces  
STRAIN: ALK04237  
FEATURE:  
NAME/KEY: exon  
LOCATION: 33..115  
OTHER INFORMATION: /codon\_start= 33  
OTHER INFORMATION: /product= "20K-cellulase"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 187..435  
OTHER INFORMATION: /product= "20K-cellulase"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 506..881  
OTHER INFORMATION: /product= "20K-cellulase"  
US-08-841-636A-30  
Query Match 30.1%; Score 353.6; DB 8; Length 936;

Best Local Similarity 72.8%; Pred. No. 42e-73;  
Matches 505; Conservative 0; Mismatches 119; Indels 70; Gaps 1;  
QY 145 CGAGTACTGGAAGCTGTCGAAGCCGTCGCTTGGCCCGGAAAGCCGCTGAGCC 204  
Db 183 CCAGTACTGGAAGCTGTCGAAGCCGTCGCTTGGCCCGGAAAGCCGCTGAGCC 242  
QY 205 AACCGCTACGCGTGTGATGCAACTTCCAGGCGCTGTCCGATCAATTCATTCGCG 264  
Db 243 AGCCGCTACGCGTGTGATGCAACTTCCAGGCGCTGTCCGATCAATTCATTCGCG 302  
QY 265 GTCGACAGGCGGCTGCGGCTTCTACTCTGCGCCACAGACTCTCTGGCGGATGAAGCA 324  
Db 303 GTCGACAGGCGGCGCCGCTTCTGTGTGCGCCACAGACTCTCTGGCGGATGAAGCA 362  
QY 325 ATCTGCTTACGCTTTCGCGGACAGACTGCGCGGCTTCCGATCTGTGTGCT 384  
Db 363 ACCTGCTTACGCTTTCGCGGACAGACTGCGCGGCTTCCGATCTGTGTGCT 422  
QY 385 GCGCTGCTACGCG----- 398  
Db 423 GTGCTGCTACGCGTGTGATGTGCTTGGGCGCAAGCTGCTGATTCGGAATTCAAGCAC 482  
QY 399 -----CTCACCTTCACTTCCGCTCCGTCGCGGCAAGCA 434  
Db 483 TGACCGAGGACCGGCTGCGGCTGCTGACCTTTCATTCGCGGCTGCGGCAAGAC 542  
QY 435 ATGTGTGTGATGATCAAGCAAGCTGCGGCTGCGGATGATTCGATTCGCTGCTGCT 494  
Db 543 ATGTGTGTGATGATCAAGCAAGCTGCGGCTGCGGATGATTCGCTGCTGCTGCT 602  
QY 495 ATGCGCGGCGGCGGCGGCTGCTTCAAGCGCTGCGGCTGCGGCTGCGGCGGCTGCT 554  
Db 603 ATGCGCGGCGGCGGCGGCTGCTTCAAGCGCTGCGGCTGCGGCTGCGGCGGCTGCT 662  
QY 555 GCGGCTCAATACGCGGCTTGTGTCGCGGACAGGCTGCTTCCGCGCGGCTGCT 614  
Db 663 GCGGCTCAATACGCGGCTTGTGTCGCGGACAGGCTGCTTCCGCGCGGCTGCT 722  
QY 615 AACCGCGCTGCGGCTGCTTGTGTCGCGGACAGGCTGCTTCCGCGCGGCTGCT 674  
Db 723 AACCGCGCTGCGGCTGCTTGTGTCGCGGACAGGCTGCTTCCGCGCGGCTGCT 782  
QY 675 TTCCAGGAGGCTGCGGCTGCTTCCGCGGACAGGCTGCTTCCGCGCGGCTGCT 734  
Db 783 TTCCAGGAGGCTGCGGCTGCTTCCGCGGACAGGCTGCTTCCGCGCGGCTGCT 842  
QY 735 GACTCCAGCTTCCGCGCTTCAACCCGCCCAAGCG 768  
Db 843 GAGCGGCGCTTCCGCGCTTCAAGGCCGCCCAAGCG 876  
RESULT 10  
US-10-007-521-3  
Sequence 3, Application US/10007521  
Publication No. US20030054539A1  
GENERAL INFORMATION:  
APPLICANT: Schulein, Martin  
Andersen, Lene N.  
Lassen, Soren F.  
Kauppinen, Markus S.  
Lange, Lene  
Nielsen, Ruby I.  
Ihara, Michiko  
Takagi, Shinobu  
TITLE OF INVENTION: No. US20030054539A1 Endoglucanases  
NUMBER OF SEQUENCES: 109  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. US20030054539A1 No. US20030054539A1disk of No. US200300C  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America





FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..924  
SEQUENCE DESCRIPTION: SEQ ID NO: 5  
US-10-007-521-5

Query Match 28.5%; Score 334.8; DB 15; Length 927;  
Best Local Similarity 64.9%; Pred. No. 1..1e-68;  
Matches 574; Conservative 0; Mismatches 277; Indels 33; Gaps 4;

```
QY 109 CTGGTCCGCTCCGCGGCAAGTGGAGTGGCCAGTCCAGAGATTAAGTGGAGCTGTGCAAG 167
DB 43 CTGGCCCTGAGACAGCTCTCGGGCATCGGCGAGAGACAGCCGGATTAAGTGGAGCTGTGCAAG 102
QY 168 CGGTGTCGCTTGGTCCGGGAGAGCGCGCTGACGCAACCGGTCTAAGCGGTGCGATGCC 227
DB 103 CGAGAGTCCGCTGCGCGGAGAGGCGCCCTGTCTCTCGGTGAGAGCGCTGCGAGCAAGAAC 162
QY 228 AACTTCAGAGCGCTGTCCGAGCTTCAATATGTCAGTCCGAGCTGCAAC---GGCGGCTCGGCG 284
DB 163 GACAAACCGCTCAAGAGAGGCGGCTCCAGCCCGCTCGGCTCGAGCGGAGCGGCGGCGCGCC 222
QY 285 TACTCTGCGCGGCAAGCAATCTCTGCGGCGGTGAACGACAACTCTGCGCTTAAGGCTTCCGC 344
DB 223 TACATGTGCTCTCCAGAGCGCTTGGGCGGTACGACGAGCTGTGTGTATGCGCTGAGCG 282
QY 345 GCGAGAGCATCGCGCGGCGGCGGTCCGAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 404
DB 281 GCGGTCAAGCTCGCGGAGGCTCTCGAGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 342
QY 405 TTCACTTCGCGGTCCGCTCGCGGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 464
DB 343 TTCAACAGCGGCGCGGCTCGCGGCGGAGAGATGATGATGATGATGATGATGATGATGATGATG 402
QY 465 GACCTGGGAAATTAACAGTGTGATATGCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 524
DB 403 GACCTGGGCGAACAACCTTTGACCTGAGCATCCCGGTGCGGCTGTGTGTGTGTGTGTGTGTGT 462
QY 525 GCGTGCAGCTGCGAGTGTGGGCGGCTGCC-----CGGCGTCAATAAGCGGCGGCAATT 575
DB 463 GCGTGCAGCTGCGAGTGTGGGCGGCTGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 522
QY 576 TGGTGGCGGCGAGTGTGATTTCTTCTCCGCGGCGGCTCAAGCCCGGCTGCGAGTGTGCGG 635
DB 523 CATTTCAAGAGAGAGTGTGATTTCTTCCGAGGCGCTCAAGCCCGGCTGCGAGTGTGCGG 582
QY 636 TTTGATGCTTCCAGAGCGCGGAGCAACCGGAGTTCAGTTCAGAGTGTGAGTGTGCGGCGG 695
DB 583 TTTGATGCTTCCAGAGCGCGGAGCAACCGGAGTTCAGTTCAGAGTGTGAGTGTGCGGCGG 642
QY 696 GCGGAGATGCTTGGCGGCTCGGCGTGCAGAGCGGAGCTTCAGGCTTCCCGGCTTTC 755
DB 643 TGGAGCTCAAGTCAAGAGGCGGCTGCTCCGATAGAGAGCGGCAATTCTTCCGCGGTC 702
QY 756 ACCCGCCCAAGCGGTGCGCAACG-----TGGCAACCGGAGAGCGCCAG 797
DB 703 CAGATCCCTCCAGAGAGCAACAGCTCTCGGCTCAACAGCTTACAGCAACAGCAACAG 762
QY 798 TGGAGTGGCTTGGGTGGGCGAAG---GTCTCCGCGGCGGCGGAGTGTGCGGCTGCAAGCT 854
DB 763 TCCAGCTTCCAGACCTTCAGGCGCGCGGCGGAGTTCAGGCTTCCAGGCGGCTGCAAGCT 822
QY 855 CAGAAAGTGGCTCAGTGGCGGTGAGTGGCTTCAAGCGGATGCAACCTGTGTCTGTGGC 914
DB 823 GAGAGTGGCTCAGTGGCGGTGAGTGGCTTCAAGCGGATGCAACCTGTGTCTGTGGC 882
QY 915 ACCAGCTGCGAAGTGAAGAGTACTTACTGTGAGTGTCTTGA 958
DB 883 AGCACTTGAAGAGATTAAGTACTGTGATCAATCAAGTCTGTGA 926
```

RESULT 12  
US-09-735-787-3

```
? Sequence 3, Application US/09735787
? Patent No. US20010036910A1
? GENERAL INFORMATION:
? APPLICANT: Rasmussen, Grethe
? Mikkelson, Jan Moller
? Schuelin, Martin
? Patkar, Shankant A.
? Hagen, Fred
? TITLE OR INVENTION: A Cellulase Preparation Comprising an
? Endoglucanase Enzyme
? NUMBER OF SEQUENCES: 33
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: No. US20010036910A10 No. US20010036910A1disk of No. US200100C
? STREET: 405 Lexington Avenue, 64th Floor
? CITY: New York
? STATE: New York
? COUNTRY: United States of America
? ZIP: 10174-6401
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/735,787
? FILING DATE: 13-Dec-2000
? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 09/189,028
? FILING DATE: <Unknown>
? ATTORNEY/AGENT INFORMATION:
? NAME: Lambiris, Elias J.
? REGISTRATION NUMBER: 33,728
? TELEPHONE: 212-867-0123
? TELEFAX: 212-878-9655
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1473 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: Pseudomonas oxydans
? STRAIN: DSM 2672
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 97..1224
? SEQUENCE DESCRIPTION: SEQ ID NO: 3:
? US-09-735-787-3
```

Query Match 26.8%; Score 315.2; DB 9; Length 1473;  
Best Local Similarity 68.5%; Pred. No. 5.1e-64;  
Matches 466; Conservative 0; Mismatches 208; Indels 6; Gaps 2;

```
QY 83 CAGAACCTTGGCGGCTGCACTTCTGTGTGCTCCGCGGCGGCAAGTGGAGTGGCGGCAAGT 142
DB 108 CACTCTTTCGCGGCTGCGCGGCTCTGTGCGGAGTGTGCTTCTGAGAGCGGTGCTC 167
QY 143 CAGCAGATATGAGAGTGTGCAAGCGGCTGTGCGCTTGGCGCGGAGAGCGCGGCTGCAAG 202
DB 168 TACTGATATGAGAGTGTGCAAGCGGCTTCTGTGTGAGAGGAGAGCGGTGCTGCAAG 227
QY 203 CCAACGCTTACGCGGTGCGAGTCCAGCTTCAAGCGGCTGTGCACTTCAATGTCCAGT 262
DB 228 CGCCCTGCTTAACTGTGATGAAGAGAGCAACCCCAATTTCAACACCAATGTCTGCA 287
QY 263 GGGCTGCAAC---CGGCGGCTGCGGCTTCTGCTGTGCGGAGCAAGACTTCCCTGGGCGGAG 319
```

Db 288 CGGTTGAGAGGAGTGGTTCCTGCTTATGCTTGACCACTACTCTCCCTGGGCTGCA 347  
QY 320 CGAGATTCCTGCTTACGGCTGCGCGGCAAGAGATTCGCGGGGTCCGATCTCTG 379  
Db 348 CGATGAGCTTGGCTTACGGTTCCTGCTTACCAAGATTCCTCGGTGGCTCCGAGGCGAGCTG 407  
QY 380 GTGCTGCGCTGCTTACGGGCTGACCTTACCTTCCGCTGCGGTGGCGGCAAGCATGGT 439  
Db 408 GTGCTGCTTGGCTTATGCTTTCCTTACCACTGAGCCCTGCTGAGGAGGAGAGATGAT 467  
QY 440 GGTGAGTCAACGAGCACTGGCGGCACTGGAGATTAACAGATTGATCGCATGGC 499  
Db 468 GTGCTGCTTACCAAGCACTGGAGGATTCCTGGGCAACCACTTGTATGATGATGTC 527  
QY 500 CGGCGCGCGCTGGGCTTCTTCAAGGCTGCAAGCTCGCATTTGCGC--GACCTCCCGG 556  
Db 528 CGGCGGTGGTGGTATCTTCAAGGCTGCAAGCTTCAAGTTCGCAAGGCTCTCGCGG 587  
QY 557 CGCTCAATACGAGCGGCACTTTCGTGGCGGACAGTGGATTCCTTCCCGCGCGCTCA 616  
Db 588 TGCCAGTACGCGGCTATCTCTCCCGAAGGATGTATAGCTTACCCGAGCTTCTCA 647  
QY 617 GCCCGCTGCGAGTGGCGGCTTTCAGTTCAGAACGCGCAACACCGAGCTTCACTG 676  
Db 648 GAGCGGTTCGCACTGGGATTCGACTGGTTCGAGAAAGCGCAACCTGACTTCACTT 707  
QY 677 CGAGAGGTGAGTGGCGCGCGGAGATGTTGCCGCTCGGCTGCAAGGCGCAAGCA 736  
Db 708 TGAGCAGGTTCAGTGGCGCGGAGCTCTCTGAGATCAGTGAATGAGTGAAGCGTATGACGA 767  
QY 737 CTCAGACTTCCCGCTTCTCA 756  
Db 768 CTCAGACTTCCCGCTTCA 787

RESULT 13  
US-10-138-870-3  
Sequence 3, Application US/10138870  
Publication No. US20030119167A1  
GENERAL INFORMATION:  
APPLICANT: Rasmussen, Grethe  
Mikkelsen, Jan Moller  
Schulein, Martin  
Pakkar, Shankant A.  
Hagen, Fred  
TITLE OF INVENTION: A Cellulase Preparation Comprising an  
Endoglucanase Enzyme  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. US20030119167A1o No. US20030119167A1disk of No. US20030119167A1  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/138,870  
FILING DATE: 03-May-2002  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/735,787  
FILING DATE: 13-Dec-2000  
APPLICATION NUMBER: 09/189,028  
FILING DATE: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 3469, 214-US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1473 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Fusarium oxysporum  
STRAIN: DSM 2672  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 97..1224  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-138-870-3

Query Match 26.8%; Score 315.2; DB 15; Length 1473;  
Best Local Similarity 68.5%; Pred. No. 5,1e-64;  
Matches 466; Conservative 0; Mismatches 208; Indels 6; Gaps 2;

QY 83 CACAACCTGGCGGCTGCACTTCTCTGTCGCTTCGCGGCGAGTGGAGGCGAGTC 142  
Db 108 CACTCTTCTGCGCTGCGGCGGCTCTGCGCTGAGTGTCTTCTGAGAGCGTCACTC 167  
QY 143 CACGAGATCTGAGATGCTGCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 202  
Db 168 TACTGATCTGAGATGCTGCAAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 227  
QY 203 CCAACCGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 262  
Db 228 CGCCCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 287  
QY 263 GGGCGGCA---CGGCGCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319  
Db 288 CGGTTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 347  
QY 320 CGAGATCTGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTAC 379  
Db 348 CGAGATCTGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTAC 407  
QY 380 GTGCTGCGCTGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTAC 439  
Db 408 GTGCTGCGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTAC 467  
QY 440 GGTGAGTCAACGAGCACTGGCGGCACTGGAGATTAACAGATTGATCGCATGGC 499  
Db 468 GTGCTGCTTACCAAGCACTGGAGGATTCCTGGGCAACCACTTGTATGATGATGTC 527  
QY 500 CGGCGCGCGCTGGGCTTCTTCAAGGCTGCAAGCTCGCATTTGCGC--GACCTCCCGG 556  
Db 528 CGGCGGTGGTGGTATCTTCAAGGCTGCAAGCTTCAAGTTCGCAAGGCTCTCGCGG 587  
QY 557 CGCTCAATACGAGCGGCACTTTCGTGGCGGACAGTGGATTCCTTCCCGCGCGCTCA 616  
Db 588 TGCCAGTACGAGCGGCTATCTTCCCGAAGGATTAAGCTTACCCGAGCTTCTCA 647  
QY 617 GCCCGCTGCGAGTGGCGGCTTTCAGTTCAGAACGCGCAACACCGAGCTTCACTG 676  
Db 648 GAGCGGTTCGCACTGGGATTCGACTGGTTCGAGAAAGCGCAACCTGACTTCACTT 707  
QY 677 CGAGAGGTGAGTGGCGCGGAGATGTTGCCGCTCGGCTGCAAGGCGCAAGCA 736  
Db 708 TGAGCAGGTTCAGTGGCGCGGAGCTCTCTGAGATCAGTGAATGAGTGAAGCGTATGACGA 767  
QY 737 CTCAGACTTCCCGCTTCTCA 756  
Db 768 CTCAGACTTCCCGCTTCA 787



SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/007,521

FILING DATE: 10-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/651,136

FILING DATE: 21-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4366, 200-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 1132 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 42..971

SEQUENCE DESCRIPTION: SEQ ID NO: 21:

US-10-007-521-21

Query Match 25.7%; Score 302; DB 15; Length 1132;

Best Local Similarity 67.4%; Pred. No. 5,9e-61;

Matches 441; Conservative 0; Mismatches 210; Indels 3; Gaps 1;

QY 111 GTGCGCTCCGCGCCAGTGGAGTGGCACTGTCACAGATATCTGGACTGTGCAAGCCG 170  
DB 81 GTGCGCGCCAGTGGTCTGTGGCAAGCGCCGACCACTACTGGACTGTGCAAGCCG 140  
QY 171 TGTGCGCTTGGCCCGGAGAGCCCGGTCAAGCAACGGGTCTAAGCGTGGTGGTGGCAAC 230  
DB 141 TCTTCCGGGGTGGAGCAAAAGGCTCCGTCAAGCAAGCCCTGTCAAGAGTGGCTAGGAG 200  
QY 231 TTCCAGCGCTGTCCGACTTCAATGTCAGTCCGCTGCAACGGCGCTCGGCTACTCC 290  
DB 201 AACCAACCTCTCGGCTCCAGCGCCAGAGAGCGGCTGCAATCCAAAGGCGCTGCAAG 260  
QY 291 TGGCGCGACAGACTCCCTGGCGGTGACAGACATCTGCTTACGGCTTCCCGGAGCG 350  
DB 261 TGCACGATACCAAGCGGTGGGCTGTCAAGATACCTGGCTATGCTTGGTGGCAAG 320  
QY 351 AGCATCGCGCGCGGTCCGAATCTCGGTGGTGGCGCTGTAGGGGCTCACTTCACT 410  
DB 321 GCTTTCAGTGGTGGATCGAGGCTCAGCTGGTGTGCTGTGCTGTATGCCCTTCAAGTCA 380  
QY 411 TCCGGTCCCGTCCGCGCCAGACATGGTGGTGGTCAATGCAAGCACTGGCGGCACTG 470  
DB 381 TCCGACCTCTGTGGGAAAGACATGGTGGTGGTCAATGCAAGCAACCGCGGCACTG 440  
QY 471 GGAAGTAAACAGTTGATATCGCAATGCGCGCGCGCGGTGGGCAATCTTCAAGGCTGC 530  
DB 441 AGCGGCAACACTTTGACATCTCAATGCGCGCGCGCGGTGGGCAATCTTCAAGGCTGC 500  
QY 531 AGCTGGCACTTGGCG---GCTTCCCGCGCGCTCAATAGGCGGCAATTTGTGCGCGAG 587  
DB 501 ACCCGCAATGGGCGCTTCCCGGAAACCGCTACGGGCGGCAACCAAGCGGCAAGC 560  
QY 588 CAGTGGATTCCTTCCCGCGCGCGCTCAAGCGCGCGCTGCAAGTGGCGTGGTGGTTC 647  
DB 561 CAGTGGTCCCAATTCCTTGGCGCTTGGCGCTGCAAGCGCGCTGCAAGTGGCGTGGTTC 620  
QY 648 CAGAAAGCGCAACCTGCACTTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 707  
DB 621 AACGACGCGCAACCTGCACTTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 680  
QY 708 GCCCGCTCGGCTGCAAGGCAAGCAAGCACTCCAGCTCCCGGTTCACCCCG 761

Db 681 GACCGACCGGCTTGGCGCGCTCCGATACAGGAACTATCCGCTTCCAGCC 734

Search completed: July 7, 2004, 15:01:27  
Job time : 4335 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2004, 08:03:07 ; Search time 4752 Seconds

(without alignments)  
7377.567 Million cell updates/sec

Title: US-10-007-521-11

Perfect score: 1174  
Sequence: 1 GAGCAGCAGCCCTCAGAGCTG.....AAAAAAAAAAAAAAAAAAAA 1174

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estha:\*  
2: em\_esthm:\*  
3: em\_estim:\*  
4: em\_estim:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estc1:\*  
10: gb\_estc2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estum:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pla:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vit:\*  
28: gb\_ges1:\*  
29: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	227.4	19.4	688	13	BU064446 Rtr 4 005
2	213	18.1	691	10	BS585661 RSTH6ESP6
3	207.6	17.7	589	14	CS928003 BSA1_35 G
4	138.2	11.8	668	14	CB408403 IPG004F12

5	137	11.7	667	14	CB407613 IPG028F05
6	132.4	11.3	681	14	CB408544 IPG007A08
7	130.4	11.1	450	12	BI200729 GLE05Fs_r
8	124.2	10.6	622	14	CB408447 IPG006G36
9	121.4	10.3	703	14	CD486264 CF08_1FL
10	114.4	9.7	444	12	BI190695 IPG029B01
11	111.8	9.5	593	14	CB407636 IPG029B01
12	111.6	9.5	630	14	CB407706 IPG027C04
13	109.8	9.4	446	14	CB408057 IPG013D01
14	109.3	9.3	534	14	CB408340 IPG010A02
15	109	9.3	568	14	CB407804 IPG023B05
16	107.4	9.1	501	14	CB408299 IPG009D10
17	106.4	9.1	426	12	BI187295 a111fs_r
18	105.8	9.0	663	14	CB407573 IPG027G07
19	100.2	8.5	582	14	CB407953 IPG023B04
20	99.6	8.5	515	14	CB408097 IPG013H06
21	98.2	8.4	433	14	CB408489 IPG016D03
22	98.2	8.4	452	14	CB408817 IPG017E01
23	95.4	8.1	571	14	CB407840 IPG024C11
24	95.4	8.1	673	14	CB408389 IPG004B07
25	79.2	6.7	601	14	CB407923 IPG021H08
26	78.8	6.7	385	14	CB408025 IPG011H04
27	74.4	6.3	447	14	CB407737 IPG025C09
28	73.4	6.3	515	13	CB424977 BX424977
29	72.6	6.2	909	13	BX31337 BX31337
30	72	6.1	562	14	CD682259 r148905_Y
31	71.8	6.1	1201	13	BX39673 BX39673
32	71.6	6.1	538	13	BX404314 BX404314
33	70.4	6.0	191	13	BX477475 BX477475
34	70.4	6.0	193	13	BX509937 BX509937
35	70.2	6.0	168	13	BQ152857 BQ152857
36	70	6.0	889	14	CF883951 CF883951
37	69.8	5.9	297	14	CB408884 IPG018D09
38	69.6	5.9	203	9	AI620089 ty48d08_x
39	69.4	5.9	520	14	CD683022 r129a08_Y
40	69.4	5.9	974	13	BX339244 BX339244
41	69.2	5.9	233	14	CF335596 JMT--05-F
42	68.8	5.9	493	12	BG656882 df25908_Y
43	68.4	5.8	804	13	BQ414074 BQ414074
44	68.4	5.8	441	12	BM573394 BM573394
45	68.4	5.8	596	13	BQ186959 UT-E-BU1-

## ALIGNMENTS

RESULT 1  
BU064446  
LOCUS  
DEFINITION  
Fgr 4 005\_T3 Nitrogen-starved mycelia *Gibberella zeae* cDNA, mRNA  
ACCESSION  
BU064446  
VERSION  
BU064446.1 GI:22504735  
KEYWORDS  
EST.  
SOURCE  
*Gibberella zeae*  
ORGANISM  
*Gibberella zeae*  
REFERENCE  
1 (bases 1 to 688)  
AUTHORS  
Trail, F., Xu, V.-R., San Miguel, P., Halgren, R.G. and Kistler, H.C.  
TITL  
Analysis of expressed sequence tags from *Gibberella zeae* (anamorph *Fusarium graminearum*)  
JOURNAL  
Fungal Genet. Biol. 38 (2), 187-197 (2003)  
MEDLINE  
22508120  
PUBMED  
12620255  
COMMENT  
Contact: Frances Trail  
Department of Plant Biology  
Michigan State University  
East Lansing, MI 48824, USA  
Tel: 517 432 2939  
Fax: 517 353 1926  
Email: trail@msu.edu  
Plate: 4 row: 0 column: 05.

## FEATURES

Location/Qualifiers  
1. 688

/organism="Gibberella zeae"  
/mol\_type="mRNA"

/strain="NRRL 31084"

/db\_xref="taxon:5518"

/clone\_lib="Nitrogen-starved mycelia"

/note="Vector: Uni-ZapII; Site\_1: EcoRI; Site\_2: XhoI"

## ORIGIN

Query Match

19.4%; Score 227.4; DB 13; Length 688;

Best Local Similarity 65.1%; Pred. No. 1.7e-16;  
Matches 368; Conservative 0; Mismatches 191; Indels 6; Gaps 2;

```

QY CCTGCGCGCTGCACTTCTCTGTCGCTCCGCGCGCAAGTGCAGTGCAGTGCAG 148
DB CTTGCGCGCTGTCGCGCTCTCGCGCTGAGTGCAGTGCAGTGCAGTGCAGTGCAG 183
QY ATATGCGAGCTCTGCAAGCGCTGTCGCTTGGCGCGAAGCGCGCGTCAAGCCACC 208
DB ATATGCGAGCTCTGCAAGCGCTTCTTGTCTTGGAGTGCAGTGCAGTGCAGTGCAG 243
QY GGTCTAGCGGTCGATGCGCACTTCAAGCGCGCTGTCGCACTTCAATGTCAGTGCAG 268
DB TGTCTTGAATGTCGCAAAAAGCAACCCATACCTGACGTCGACGTCGTCGTCGTCG 303
QY 269 ---CAAGCGCGCTGCGCTTCTCTGTCGCGCGCACTTCTCTGCGCGTGAACGCA 325
DB TGAAGTGTGTGTCGCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 363
QY 326 TCTGCGCTGAGCTTCCGCGCGCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 385
DB CTTGCTTGAAGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 423
QY 386 CGCGCTGTCAGCGCTCACTTCTCTGTCGCTCCGCGCGCAAGTGCAGTGCAGTGCAG 445
DB TGTCTTGAATGTCGCAAAAAGCAACCCATACCTGACGTCGACGTCGTCGTCGTCG 483
QY 446 GTCAACGAGCACTGCGCGCGCACTGCGAGTGAACGTCGTCGTCGTCGTCGTCGTCG 505
DB GTCCACCAACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 543
QY 506 CGCGTGTGCGATCTTCAAGCGCTGCGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 562
DB 544 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 603
QY 563 ATACGCGCGATTCGTCGCGCGCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 622
DB 604 GTACGCGCGATCTTCTCCGAGCAATGTGACGTCGTCGTCGTCGTCGTCGTCGTCG 663
QY 623 CTGCGAGTGCAGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 647
DB 664 TTGCCACTGCGCATTTGACTGAGTC 688

```

## RESULT 2

BE585661

LOCUS

DEFINITION BE585661 691 bp mRNA linear EST 17-AUG-2000  
EST#68SP6.D02.d2.014 KSU wheat Fusarium graminearum infected spike  
CDNA library Triticum aestivum cDNA clone EST#68SP6.D02.d2.014,  
mRNA sequence.

## ACCESSION

BE585661

## VERSION

BE585661.1

## KEYWORDS

EST.

## SOURCE

Triticum aestivum (bread wheat)

## ORGANISM

Triticum aestivum

## REFERENCE

Fellers, J.P., Li, W.L., Hill-Ambroz, K., Matthews, A. and Gill, B.S.

## AUTHORS

The structure and function of the expressed portion of the wheat

## TITLE

genomes - Kansas State University. Fusarium graminearum infected

## JOURNAL

sipike cDNA library  
Unpublished (2000)

## COMMENT

Contact: John Fellers  
US Department of Agriculture, Agriculture Research Service, Plant  
Science and Entomology Unit  
Dept. of Plant Pathology, 4006 Throckmorton Hall, Kansas State  
University, Manhattan, KS 66506, USA  
Tel: 785-532-2367  
Fax: 785-532-6167  
Email: jpf@ksu.edu

Sequence have been trimmed to remove vector sequence and low  
quality sequence with phred score less than 20  
Seq primer: SP6.

## FEATURES

Location/Qualifiers  
1. 691

/organism="Triticum aestivum"  
/mol\_type="mRNA"

/cultivar="Sumai3"

/db\_xref="taxon:4565"

/clone="EST#68SP6.D02.d2.014"

/tissue\_type="Spike"

/dev\_stage="Adult plant"

/lab\_host="B. coli UM109"

/clone\_lib="KSU wheat Fusarium graminearum infected spike  
cDNA library"

/note="Vector: pGEM-T easy; Site 1: SacI; Site 2: SpeI;  
Plants were grown in the greenhouse. Spikes were sprayed  
with Fusarium graminearum (at what stage). Total RNA, and  
poly(A) RNA were prepared from infected spikes. cDNA was  
prepared using the SmartTM PCR cDNA synthesis kit from  
Clontech. cDNA was cloned into the pGEM-T easy vector  
from Promega."

## ORIGIN

Query Match

18.1%; Score 213; DB 10; Length 691;

Best Local Similarity 64.9%; Pred. No. 6.7e-15;  
Matches 364; Conservative 0; Mismatches 190; Indels 7; Gaps 3;

```

QY CCTGCGCGCTGCACTTCTCTGTCGCTCCGCGCGCAAGTGCAGTGCAGTGCAG 148
DB CTTGCGCGCTGTCGCGCTCTCGCGCTGAGTGCAGTGCAGTGCAGTGCAGTGCAG 190
QY 89 CCTGCGCGCTGCACTTCTCTGTCGCTCCGCGCGCAAGTGCAGTGCAGTGCAG 208
DB 131 CTTGCGCGCTGTCGCGCTCTCGCGCTGAGTGCAGTGCAGTGCAGTGCAGTGCAG 250
QY 149 ATATGCGAGCTCTGCAAGCGCTTCTGTCGCGCGCAAGTGCAGTGCAGTGCAG 268
DB 191 ATATGCGAGCTCTGCAAGCGCTTCTGTCGCGCGCAAGTGCAGTGCAGTGCAG 310
QY 209 GGTCTAGCGGTCGATGCGCACTTCAAGCGCTGTCGCACTTCAATGTCAGTGCAG 325
DB 251 TGTCTTGAATGTCGCAAAAAGCAACCCATACCTGACGTCGACGTCGTCGTCGTCG 370
QY 269 ---CAAGCGCGCTGCGCTTCTCTGTCGCGCGCAAGTGCAGTGCAGTGCAGTGCAG 385
DB 311 TGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 430
QY 326 TCTGCGCTTACGAGCTTCCGCGCGCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 445
DB 371 CTTGCTTACGAGTTCACGCTTACCAAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 490
QY 386 CGCGTGTGCGATCTTCAAGCGCTGCGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 505
DB 431 TGTCTTGAATGTCGCAAAAAGCAACCCATACCTGACGTCGACGTCGTCGTCGTCGTCG 550
QY 446 GTCAACGAGCACTGCGCGCGCACTGCGAGTGAACGTCGTCGTCGTCGTCGTCGTCGTCG 562
DB 491 GTCAACCAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 610
QY 506 CGCGTGTGCGATCTTCAAGCGCTGCGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 622
DB 551 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 669
QY 563 ATACGCGCGATTCGTCGCGCGCAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 691
DB 611 GTACGCGCGATCTTCTCCGAGCAATGTGACGTCGTCGTCGTCGTCGTCGTCGTCGTCG

```









